Working paper

International Growth Centre

Reshaping Institutions

Evidence on External Aid and Local Collective Action

Katherine Casey
Rachel Glennerster
Edward Miguel

April 2011

When citing this paper, please use the title and the following reference number:
S-1002-SLE-1

DIRECTED BY LSE UNIVERSITY OF OXFORD

FUNDED BY UKaid from the British people

Reshaping Institutions: Evidence on External Aid and Local Collective Action

Katherine Casey
Brown University

Rachel Glennerster
Abdul Latif Jameel Poverty Action Lab

Edward Miguel
University of California, Berkeley and NBER

FIRST DRAFT: APRIL 2011
THIS DRAFT: 26 APRIL 2011

Abstract: Although institutions are believed to be key determinants of economic performance, there is limited evidence on how they can be successfully reformed. The most popular strategy to improve local institutions in developing countries is "community driven development" (CDD). This paper estimates the impact of a CDD program in post-war Sierra Leone using a randomized experiment and novel outcome measures. We find positive short-run effects on local public goods provision, but no sustained impacts on fund-raising, decision-making processes, or the involvement of marginalized groups (like women) in local affairs, indicating that CDD was ineffective at durably reshaping local institutions.

Acknowledgements: We wish to thank the GoBifo Project staff—Minkahil Bangura, Kury Cobham, John Lebbie, Dan Owen and Sullay Sesay—and the Institutional Reform and Capacity Building Project (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, NEUDC, WGAPE, and U.C. Berkeley have provided helpful comments. We gratefully acknowledge financial support from the GoBifo Project, the IRCBP, the World Bank Development Impact Evaluation (DIME) 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). All errors remain our own.

Corresponding author: Edward Miguel (emiguel@econ.berkeley.edu)

Experience demonstrates that by directly relying on poor people to drive development activities, CDD [community driven development] has the potential to make poverty reduction efforts more responsive to demands, more inclusive, more sustainable, and more cost-effective than traditional centrally led programs…achieving immediate and lasting results at the grassroots level. – Dongier et al. (2003)

1. Introduction

Many scholars have argued that the accountability and inclusiveness of government institutions are key determinants of economic performance. In particular, institutions that are egalitarian and protect individual rights have been tied to better economic outcomes in India (Banerjee and Iyer 2004), Brazil and the United States (Engerman and Sokoloff 1997), and former European colonies (Acemoglu, Johnson and Robinson 2001). However, there is no consensus on the specific policy reforms that will successfully engender better functioning institutions, or on whether it is possible, or even desirable, for external actors like foreign aid donors to attempt to reshape local power dynamics in less developed countries. This debate has played out vigorously in discussions on aid policy reform: while some scholars argue that large infusions of foreign aid can themselves help build stronger institutions (Sachs 2005), others assert that historically rooted local institutions and social norms are difficult to understand, let alone transform (Easterly 2006), and that attempts by outsiders to create "better" institutions will be futile.

Among foreign aid donors, non-governmental organizations (NGOs) and governments in less developed countries today, the most popular strategy to promote the accountability, competence and inclusiveness of local institutions lies in "community driven development" (CDD). Over nine percent of total World Bank lending supports CDD projects, placing annual investment in the billions of U.S. dollars (World Bank 2007). Typical CDD interventions combine block grants for local public projects with intensive training that aims to both streamline and catalyze collective action. In particular, CDD attempts to bolster local *coordination* – for example, by setting up village development committees (VDC) and plans – and to enhance *participation* and inclusion, by requiring women and members of other marginalized groups to attend project meetings and hold leadership positions. Moreover, as arguments in favor of the CDD approach are related to those behind fiscal decentralization, it is often used to provide "bottom up" support for broader decentralization reforms in practice.

Yet while CDD advocates promise a long and varied list of benefits – ranging from more cost-effective construction of local infrastructure, to a closer match between project choice and

village needs, to the weakening of authoritarian village institutions (as illustrated in the opening quotation) – critics hold concomitant concerns that CDD participation requirements serve as a regressive tax, widening political participation will clog up rather than expedite decision-making (Olson 1982), external resources may attract new leaders and crowd out the more disadvantaged (Gugerty and Kremer 2008), and that these additional resources will be captured by elites if the program is unable to change the nature of *de facto* political power (Bardhan 2002). While researchers have begun to explore these claims, few studies provide rigorous evidence on the real-world impacts of community driven development projects (Mansuri and Rao 2004).

This paper studies a large-scale randomized CDD project in Sierra Leone, a country that provides a challenging yet compelling environment in which to assess the impacts of the CDD approach. The country has a dual system of governance in which the national state apparatus based in the capital Freetown runs in parallel to the "traditional" local chieftaincy system, neither of which has historically been particularly democratic or inclusive (Mamdani 1996). Regarding the former, authoritarian leaders in the 1970's and 1980's enriched themselves through illicit diamond deals while providing woefully inadequate public services, leading to protracted institutional decline (Reno 1995). President Siaka Stevens dismantled democratic institutions entirely, initially by abolishing elected district governments in 1972, and ultimately declaring the country a one-party state in 1978. One-party rule continued until the 1992 coup that roughly coincided with the start of the civil war (which ran from 1991 to 2002).

As background on the traditional system, the 149 paramount chiefs come from hereditary "ruling houses;" they serve for life once appointed or elected; and exert considerable control over resource allocation, including land and labor, as well as the local court system that reigns outside the capital. Dominated by male elders, this system has continued to the present day to largely exclude women (who are not even eligible to serve as chiefs in much of the country) and young men from decision-making. Growing frustration with government incompetence and corruption, grievances against heavy-handed chiefs, and the exclusion of women and young men from politics are seen as destabilizing factors that contributed to war (Richards 1996, Keen 2003).

Emerging from the war with dismal standards of living, health and education that placed Sierra Leone at the very bottom of the United Nations Development Program Human Development Index (United Nations 2004), the government and its donor partners launched major institutional reforms to restore multi-party democracy and stimulate economic growth. In

this context, CDD's emphasis on rebuilding village institutions while simultaneously funding basic public services seemed appropriate and provided grassroots support to the 19 newly re-established district councils. The program we study, "GoBifo" (or "Move Forward" in Krio, Sierra Leone's *lingua franca*), was a government project funded predominantly by the World Bank. GoBifo provided both what we in this paper call "hardware" and "software" support to rural communities. The hardware included block grants of $4,667, or roughly $100 per household, for constructing local public goods and sponsoring trade skills training and small business start-up capital. Akin to community organizers in the U.S., GoBifo program facilitators also provided technical assistance that promoted democratic decision-making, the participation of socially marginalized women and youth in local politics, and transparent budgeting practices. To formally link GoBifo activities into the broader decentralization reform taking place, VDC's were required to submit their village development plans to the appropriate Ward Development Committee (WDC) for review, endorsement and onward transmission to the district council for approval (GoBifo Project 2007).[1] While the objective of making local government institutions more participatory aimed to address some of the perceived root causes of the civil war, GoBifo's design is similar to many other CDD projects in non-post-conflict societies.

This paper assesses the extent to which GoBifo achieved its goals of improving local governance in rural Sierra Leone communities, and in so doing makes five contributions to the literature on aid, institutions and economic development.

We first develop a theoretical framework for understanding the mechanisms through which CDD programs might impact public goods and collective action both during and after the program, and use the model to guide the analysis and interpretation of our empirical results. As most existing evaluations of CDD projects have been relatively atheoretical, we believe this model could be useful in interpreting the findings of other studies as well.

The second important aspect is the randomized experimental research design, which produces rigorous evidence on causal impacts in a relatively large study sample of 236 villages and 2,832 households. Third, the extended timeframe of our study over four years (2005-2009) allows us to assess longer run impacts on institutional outcomes than is typically possible. While four years may be short in comparison to the lifetimes over which current institutions developed,

[1] The WDCs are the lowest formal government administrative unit, covering around 10,000 citizens on average, and the elected district councilor representing the ward serves as the WDC chair. While the project we study also operated at the ward level, only the village-level intervention was randomly assigned and is thus our focus.

it is not short in comparison to the time scales of most CDD or other externally funded projects.

Fourth, the research and project teams agreed to a set of hypotheses regarding the likely areas of program impacts in 2005 before the intervention began. As the project came to a close in 2009, we 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 analyzing the follow-up data (see supplementary web Appendix A). 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.[2]

Lastly, we combine rich household survey data with novel "structured community activities" (SCAs) that introduce three concrete, real-world scenarios that allow us to observe and measure how communities: (i) respond to a matching grant opportunity to purchase building materials at a subsidized price; (ii) make a communal decision between two comparable alternatives; and (iii) allocate a valuable asset (provided for free) among community members. We feel that these SCAs capture 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. The fact that these activities were carried out *after* the GoBifo program (and its financial resources) had ended allows us to isolate any persistent impacts on collective action and institutional performance generated by the program. We are unaware of other studies that have used this type of SCAs in practice, and believe they may be useful tools for other researchers interested in unobtrusively measuring local collective action patterns.

The analysis explores an exceptionally wide range of outcome measures, which we divide into three broad groups: project implementation outcomes (which we call "family A"), "hardware" or local public infrastructure outcomes (family B), and "software" or institutional and collective action outcomes (family C). We find that the GoBifo project was well implemented: it successfully established the village-level organizations and tools to manage development projects in nearly all cases, and provided communities with the financing to implement them (family A). The distribution of project benefits within communities was equitable for the most part, and the leakage of project resources was minimal. We also find that

[2] See Simes (1986), Horton and Smith (1999), and the NIH-supported web registry clinicaltrials.gov.

CDD had immediate impacts on local public goods infrastructure "hardware." Treatment villages have a larger stock of higher quality local public goods than control areas, and there is more market activity in treatment communities, including the presence of more traders and items for sale, suggesting short-run economic gains (family B).

There is, however, no evidence that the program led to fundamental changes in the "software" of collective action – namely, local fundraising capacity, decision-making processes, or even social attitudes and norms. As an example, despite the new experiences many women in treatment villages gained by participating in GoBifo activities, they were no more likely to voice an opinion in community meetings after the project ended or to play a leading decision-making role (along various other metrics). Similarly, the establishment of a democratically elected village development committee that carried out multiple projects did not lead treatment villages to be any more successful at raising funds in response to the matching grant opportunity. These patterns, and the lack of significant effects across many other outcomes, indicate that the program did not reshape village institutions or improve collective action beyond the activities directly stipulated by the project itself. The time horizon of the research over four years suggests that these results cannot be dismissed out of hand as the result of a short term study.

By emphasizing this project's failure to transform local institutions, we do not suggest that reforming institutions is impossible. Indeed, our results provide an interesting contrast to other recent studies that find impacts, albeit in different contexts. Beamen et al. (2009) show that requiring village councils in India to elect a woman leader reduced statistical discrimination against female candidates within a few years. Indian political reservations give representatives of historically excluded groups real power over resources as part of a formal state body (the *panchayat*). By contrast, CDD takes a more indirect approach to *de jure* reforms—nudging communities to become more democratic and inclusive without explicitly attempting to weaken elites—and may in reality not change the identity of *de facto* local power holders (Acemoglu and Robinson 2008). Perhaps because sidelining the chiefs was not a program goal, male elders and chiefdom officials retained just as much control over village development committees in GoBifo communities as they held over comparable organizing bodies in control areas, despite requirements that women attend meetings and play a role in project management.

The finding that a well implemented project with beneficial public goods and economic impacts did not trigger broader spillover effects on institutions and norms resonates with the

mixed results seen in the emerging empirical literature on CDD programs.[3] While their full academic paper is not yet available, Fearon, Humphreys and Weinstein (2009) concurrently conducted a randomized evaluation of a community driven post-war reconstruction project in Liberia in 83 communities. 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 of our findings. Yet 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 internally displaced persons (IDPs), while there are few remaining IDPs in our research sites. More speculatively, the Liberia program operated in what was the "epicenter" of the latter years of that country's civil war, and thus may have faced more disruption to local institutions than the Sierra Leone program did. Attempts to create new institutions and norms where formal structures have broken down may encounter less resistance than efforts to persuade existing authorities – such as the chiefs in our study– to adopt new practices. Crucially, though, neither our study nor Fearon et al. (2009) finds compelling evidence of program 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.

The rest of the paper is structured as follows. Section 2 lays out a simple theoretical framework of the local collective action decision in the context of a community driven development program. Section 3 discusses the GoBifo intervention, research design, econometric specifications and the *ex ante* analysis plan in greater depth. Section 4 discusses the empirical results and accompanying robustness checks, and the final section concludes.

[3]In the Philippines, Labonne and Chase (2008) find that CDD increased participation in village assemblies and interaction between residents and village leaders but did not initiate broader social change, and in fact, may have crowded out other avenues for collective action. Voss (2008) finds mixed impacts of the Kecamatan Development Program (KDP) in Indonesia on household welfare and access to services. 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 mobilization on public service providers similarly finds mixed results with strong positive effects seen for healthcare in Uganda (Bjorkman and Svensson 2009) but no effect on education in India (Banerjee et al. 2010).

2. Collective Action and Community Driven Development

We lay out a stylized local collective action framework that clarifies how an external CDD intervention might change local decision making and institutions, and derive implications that then structure our empirical analysis. In the model, a social planner determines the optimal investment in local public goods and sets a corresponding tax schedule, which is implemented with perfect compliance. Individual residents then decide whether or not to voluntarily participate in the planning and implementation of the public goods projects, taking their individual tax burden as given. We feel this framework is a reasonable approximation to the context of rural Sierra Leone (and similar societies with strong headmen), where the traditional village chief has the authority to levy fines and collect taxes to provide basic public goods, but there is variation in how involved residents are in decision making and implementation.

We define three time periods: $t = 0$ denotes the pre-program period, $t = 1$ is the program implementation phase, and $t = 2$ is post-program. Our data correspond to these three time periods: the baseline survey was fielded in t_0, the first follow-up survey captured activities that had been completed during the CDD intervention and launched the structured community activities, and the second follow-up survey explored what happened with the SCAs after the project had finished. Studying the post-program period allows us to evaluate the persistence and "sustainability" of CDD impacts.

First consider the individual's decision of whether to contribute time and voluntary labor to the planning and implementation of local public goods. While these decisions are taken in a decentralized fashion, they will aggregate in a way that affects the costs of public goods provision facing the social planner. The fact that individuals ignore the aggregate effect of their voluntary labor captures the classic externality feature of collective action, and implies that even with perfect tax compliance, the planner will still not be able to achieve the first-best level of local public goods provision. The utility maximization problem for individual *i* at time *t* is:

$$max_{\omega_t}\, u_{it} = V(g_t) + x_{it} + b_{it} * \omega_{it} \tag{1}$$

subject to the budget constraint:

$$y_{it} = x_{it} + k * \omega_{it} + \tau_{it} \tag{2}$$

where $V(\cdot)$ is a concave function capturing utility derived from consumption of the current stock of public goods g_t, where we assume for simplicity that $V(\cdot)$ is the same for all residents; x_{it} is private consumption; b_{it} is the individual's psychic or social benefit of participating in collective

action, which captures the intrinsic value of civic involvement;[4] and ω_{it} is an indicator variable that equals one if the individual participates and zero otherwise. Given historical legacies of exclusion, we assume that while some women and youth may derive positive utility from participation they face additional social costs of speaking up and thus, on average, their net benefits of civic participation are lower than for the traditional elder male elites.

The individual's budget constraint is determined by exogenous potential income y_{it}, which we define as disposable income beyond a subsistence threshold; the cost of project participation $k > 0$, which reflects the opportunity cost of time spent engaging in public goods provision instead of wage-earning activities; and the tax τ_{it}, which is set by the social planner. For simplicity, assume a tax proportional to net income, $\tau_{it} = \tau_t(y_{it} - k * \omega_{it})$, where $\tau_t \in [0,1]$ is the tax rate. The first order conditions imply that the individual chooses to participate in collective action if and only if the net benefits are nonnegative: $b_{it} - k(1 - \tau_t) \geq 0$.

Note that our model concerns only the quantity of public goods and not their type. An alternative way to conceptualize the problem would be to assume that residents have heterogeneous preferences over the type of good—i.e., a primary school versus a latrine—and that the strength of these preferences drives the choice to participate. For example, Alesina, Baqir and Easterly (1999) show that communities with more polarized preferences agree to lower tax contributions and thus fewer public goods. Similarly, Osborne, Rosenthal and Turner (2000) find that when participation is costly only a subset of residents with the most extreme views will attend meetings to determine public choice on an issue. We do not pursue this approach as we find remarkably muted differences in public spending preferences across social groups in the baseline data.[5]

Next consider the social planner's local public goods investment decision for the current

[4] We do not separate out being actively involved in decision making and involvement in provision (voluntary labor) as one usually necessitates the other as anyone who has spoken up in a meeting of a voluntary group can attest.

[5] For example, we see few differences in the priorities expressed by men as compared to women, and youths as compared to their elders, in response to the baseline survey question "If your community was given 5 million (5,000,000) Leones (US$1,667), what do you think the community should spend it on first?" Specifically, 25.1% of women versus 28.3% of men cited education projects as their first choice (along with 27.1% of youths versus 26.1% of non-youths); 13.0% versus 11.0% cited water and sanitation (12.0% versus 11.9%); 10.4% versus 10.3% cited health (10.6% versus 10.5%); and 9.2% versus 11.2% cited agriculture (10.7% versus 9.9%). Simple t-tests suggest that none of these differences are significant at 95% confidence. Moreover, Glennerster, Miguel and Rothenberg (2010) find no evidence that ethnic diversity, which could proxy for polarized preferences, inhibits local public goods provision in Sierra Leone. Although we are unable to test this hypothesis here, if heterogeneous preferences are more pronounced in other settings, the consensus building process emphasized by CDD facilitators might conceivably narrow the differences in preferences across groups and thereby trigger greater public goods provision.

time period, $q_t \geq 0$, given the stock of public goods inherited from the previous period, g_{t-1},and assuming that the stock depreciates at rate δ from one period to the next. The planner's objective is to maximize the sum of individual utilities in period *t*:

$$\max_{q_t,\tau_t} \sum_{i=1}^{N} u_{it} = \sum_{i=1}^{N} V(\delta g_{t-1} + q_t) + (1-\tau_t)(\sum_{i=1}^{N} y_{it} - k\omega_t) + \sum_{i=1}^{N} b_{it} * \omega_{it} \quad (3)$$

subject to the budget constraint:

$$c_t(q_t) = p_t q_t + \gamma_t(\omega_t, \varphi_t) \leq \tau_t(\sum_{i=1}^{N} y_{it} - k\omega_t) \quad (4)$$

where the cost function c_t has a marginal component, $p_t q_t$, where $p_t$is the price of construction materials, as well as a fixed coordination cost of collective action γ_t, which is a function of the sum of individual participation decisions ($\omega_t \equiv \sum_{i=1}^{N} \omega_{it}$) and the capacity of local institutions, φ_t.

Following the theory motivating CDD, we assume that the fixed costs of collective action are falling in both the capacity of local institutions ($\frac{\partial \gamma_t}{\partial \varphi_t} \leq 0$) and community participation ($\frac{d\gamma_t}{d\omega_t} + \tau_t k \leq 0 \; \forall \; \tau_t$); we assess the empirical validity of these assumptions below. The latter condition would be true if, for example, greater community involvement made public goods provision easier and if more involvement in decision making created greater support for the process.[6] An alternative perspective, which we do not focus on here, is that this derivative switches sign at sufficiently high participation levels if, for example, the expression of too many opinions leads to conflict or congestion in deliberation (Olson 1982) or impairs technical decision-making (Khwaja 2004). This reversal in sign may plausibly be more evident in contexts with larger baseline heterogeneity in preferences over public goods. Importantly, even if participation has no effect on coordination costs at all, CDD advocates argue that local civic engagement carries intrinsic benefits, and therefore project participation belongs in the individual utility function and its enhancement becomes an appropriate objective for intervention.

The standard Kuhn-Tucker Lagrangean for the planner's optimization problem is:

$$NV(\delta g_{t-1} + q_t) + (1-\tau_t)(\sum_{i=1}^{N} y_{it} - k\omega_t) + \sum_{i=1}^{N} b_{it} * \omega_{it} - \lambda\Big(p_t q_t + \gamma_t(\omega_t, \varphi_t) - \tau_t(\sum_{i=1}^{N} y_{it} - k\omega_t)\Big) \quad (5)$$

[6]Olken (2010) finds that choosing local development projects by direct voting instead of representative meetings increases satisfaction with and the perceived legitimacy of the project, even though the choice process has no impact on the type of project selected. He also finds that voting increases plans to use and contribute labor to the project and beliefs about the project's fairness and value. In a laboratory setting, Dal Bó et al. (2010) show that cooperation increases more when players vote to implement a change in payoffs that facilitates coordination than when the same change is imposed exogenously, again implying that having a direct say in the decision-making process can have an effect on behavior beyond the impact of the specific policy choice *per se*.

The first order conditions imply that the planner either chooses the efficient level of investment (q_t^*) with a corresponding tax rate (τ_t^*), or zero public investment and no taxes. Given the extreme poverty and limited public services in rural Sierra Leone, assume that the marginal benefit of public goods is greater than private consumption, yielding an interior solution:

$$q_t^* = V'^{-1}\left({}^{p_t}/_{N}\right) - \delta g_{t-1} \tag{6}$$

$$\tau_t^* = {}^{\left(p_t q_t^* + \gamma_t(\omega_t, \varphi_t)\right)}\Big/_{\left(\sum_{i=1}^{N} y_{it} - k\omega_t\right)} \tag{7}$$

The planner chooses q_t^* if it is affordable, and $\overline{q_t} < q_t^*$ that exhausts the budget if it is not.

Within this framework, a textbook CDD intervention aims to have three separate impacts. First, by subsidizing the cost of construction materials, the financial grants reduce the marginal cost of public goods provision, p_t. Second, participation requirements for women and youth aim to increase the benefits of participation for these historically marginalized groups. As examples, GoBifo required that one of the three co-signatories on the community bank account be female; encouraged women and youths to manage their own projects (e.g., small business training for women); made evidence of inclusion in project implementation a prerequisite for the release of block grant funding tranches; and, as part of their internal review process, required field staff to record how many women and youth attended and spoke up in meetings. Such requirements should automatically translate into greater participation in collective activities during project implementation for these groups. Moreover, if women and young men learn-by-doing, or if their participation exerts positive demonstration effects on others, shifting social norms, this experience could trigger a persistent increase in their benefits of participation, b_{it}, sustainably raising ω_t. Third and finally, this increase in community participation, accompanied by the establishment of village development committees, plans and bank accounts (boosting φ_t), aims to reduce the fixed coordination costs of collective action, γ_t. The idea is that once an organizing body is in place and residents have reached consensus on local priorities, the next collective project should be less costly to identify and execute. As such, the original GoBifo project funding proposal emphasizes the sustainability, "durability" and broad mandate of these new structures, suggesting they will become "the focal point for development interventions" and other forms of local collective action in the future (World Bank 2004).

How these three distinct effects will alter public goods investment depends on whether the village budget constraint binds. We begin with the benchmark interior solution case, where

the budget constraint is not binding, noting that we feel this case is less empirically realistic and so limit our discussion here to a brief summary (see Appendix B for a full exposition). The social planner will chose an investment amount q_t^* (as defined in equation 6) that brings the existing stock of public goods up to the efficient level g_t^*, which is defined as the level at which the sum of marginal benefits exactly equals the marginal cost of the last unit of public investment. To maintain this steady state level, current investment exactly replenishes the loss in last period's stock due to depreciation. By reducing the *marginal* cost of public infrastructure construction (and hence the efficient level of provision), CDD grants will lead to a temporary increase in public goods during project implementation. Note, however, that the reduction in coordination costs will have no effect (beyond increasing private consumption), because lowering the *fixed* cost does not change the optimal level of provision. In the post-program period, the community optimally draws down the "artificially" high levels of public good investments (due to the temporary construction price subsidies) back to the pre-program steady state level. Thus in this unconstrained case, we expect a large subsidy of material costs to trigger a short term increase in public goods followed by a post-program contraction in investment.

Given that Sierra Leone is one of the world's poorest countries, it seems more reasonable to assume that study communities face a binding budget constraint that keeps public investment well below optimal levels. This means that there are plenty of public investments—in latrines, water wells, primary schools—whose village-wide marginal benefits exceed the marginal cost of construction, yet are simply unaffordable given the community's tax base and inability to borrow in light of pervasive financial market imperfections. Under these constraints, profitable investments become unaffordable because construction prices and/or coordination costs are prohibitively high. As with the interior solution above, here we expect the financial subsidy to increase current investment in public goods during CDD project implementation. We will now also find that reducing fixed coordination costs enhances public investment, and moreover, this effect should be evident in both the implementation and post-program periods. Compared to the first, this second constrained budget case thus better corresponds to the claims by CDD advocates that a temporary intervention can permanently improve the quality of local public goods by reducing the costs of collective action.

To see this, note that when the budget constraint binds, the planner chooses the maximum affordable investment as determined by the total cost function (in equation 4 with $\tau_t = 1$):

$$\overline{q_t} = {}^{1}/_{p_t}\left(\sum_{i=1}^{N} y_{it} - k\omega_t - \gamma_t(\omega_t, \varphi_t)\right) \quad (8)$$

The affordable investment is decreasing in both construction prices and coordination costs:

$$\frac{\partial \overline{q_t}}{\partial p_t} = {}^{-1}/_{p_t^2}\left(\sum_{i=1}^{N} y_{it} - k\omega_t - \gamma_t(\omega_t, \varphi_t)\right) < 0 \quad (9)$$

$$\frac{\partial \overline{q_t}}{\partial \gamma_t} = {}^{-1}/_{p_t} < 0 \quad (10)$$

Thus in $t = 1$, the CDD project relaxes the budget constraint by both lowering the marginal cost of materials through the grants, and decreasing the fixed coordination costs by enhancing community participation and capacity. These together imply that current investment increases ($\overline{q_1} > \overline{q_0}$). In $t = 2$, communities continue to enjoy greater public investment due to the reduced coordination costs ($\gamma_2 < \gamma_0$), however, the loss of the financial subsidy places current investment somewhere between baseline and implementation levels ($\overline{q_0} < \overline{q_2} < \overline{q_1}$).

The constrained scenario is depicted graphically in Figure 1. Note that the total investment associated with obtaining the optimal public goods stock (which is g_o^* in steady state, defined by the intersection of the marginal cost and benefit curves) is not affordable in any time period by assumption ($\sum_{i=1}^{N} y_{it} < c_t(q_t^*)$). Current investment is instead determined by the intersection of the total cost and budget lines at a corner. Assuming that the community is in steady state where current investment exactly replenishes the depreciation of last period's stock[7], the planner faces a stock of δg_0 in both t_0 and t_1. Because the $c_1(q)$ cost line has both a lower intercept and flatter slope than the baseline $c_0(q)$, investment $\overline{q_1}$ exceeds pre-program steady state investment $\overline{q_0}$. In $t = 2$, the end of the subsidy adjusts the slope of the cost line back to baseline levels, while the durable improvements in coordination maintain the lower intercept. This combination allows current investment $\overline{q_2}$ to again exceed $\overline{q_0}$, but by less than in $t = 1$.[8,9]

This framework generates three empirical predictions. First, the combination of financial subsidies and lower coordination costs should unambiguously increase public goods investment

[7] The alternative assumptions would be a current investment that exceeds or falls short of replenishment. The steadily increasing stock implied by the former would eventually lead to the unconstrained case; while the steady decrease implied by the latter would suggest a starting point near zero, with no substantive changes to our results.
[8] Whether the final stock in $t_2$exceeds that in $t_1$or not depends on the depreciation rate and the size of the subsidies.
[9]We considered possible income effects. The quasilinear utility function rules out the possibility that higher local income change local demand for public goods. A plausible alternative would be to make the cost of participating a function of income ($k_{it} = f(y_{it})$), so that as people become wealthier the opportunity cost of their time increases and makes them less willing to attend meetings. Given that the constrained case appears a better match for our empirical setting (and the fact that program grants were quite modest), the first order effects of increased income would be to shift out the budget constraint and increase the current investment in public goods. Any offsetting negative effects due to higher opportunity costs of time in this setting would likely be of second order.

during the program implementation phase. To assess this, indicators under outcome family A ("implementation") below first evaluate whether the grants were in fact delivered to villages and new institutions established on the ground. The organizational capacity (community participation) of control communities is $\varphi_0(\omega_0)$ in our model, which we expect to be weaker than in treatment communities, as captured in $\varphi_1(\omega_1)$. The measures in outcome family B regarding the stock of local public goods allow us to assess the impact of the CDD program support on public investment levels, where public goods investment in treatment (control) communities during the life of the program corresponds to $q_1(q_0)$ in the model.

Second, the model implies that establishment of durable village institutions should lead to greater investment in public goods in the post-program period, which is captured most directly by the take-up of the building materials vouchers in SCA #1, as well as several other measures in outcome family C concerning collective action beyond the direct program sphere. Post-program outcomes in the treatment villages correspond to q_2, again versus q_0 in the control group.

Third, if CDD project participation requirements for women and youth trigger a permanent enhancement in the net individual benefit of participation they experience, we should see more women and youths attending community meetings and taking part in decision-making post-program (reflecting ω_2 in the treatment group and ω_0 in control), as captured by several outcomes in the gift choice component of SCA #2 and household survey responses concerning civic engagement in non-program spheres. Moreover, enhancing participation by marginalized groups could initiate broader changes in social norms and attitudes, as captured in several additional hypotheses under outcome family C examining this community "software."

3. Research Design and Analysis

3.1 The GoBifo Project

The process of establishing new village institutions, training community members, and promoting social mobilization of marginalized groups was intense and accounted for a large part of GoBifo human and financial resources. Specifically, all project 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. After the start of project work in January 2006 and through the completion of all village-level projects in July 2009, each village thus received roughly six months of direct "facilitation" over a three and a half year period (see the detailed timeline in

Appendix C). Furthermore, while just under half of the total GoBifo budget was dedicated to village- and ward-level block grants (US$896,000 or 47%), the balance covered "capacity development" in village- and ward-level planning (US$589,732 or 30%), project management and contingencies (US$255,320 or 14%), and monitoring and evaluation (US$177,300 or 9%). Thus for every dollar spent directly on actual construction or training projects, roughly one dollar was spent on capacity-building, facilitation and oversight.

GoBifo village projects were carried out in several areas: the largest share, at 43%, was in the construction of local public goods, with 14% in community centers or sports fields, 12% in education (i.e., primary school repairs), 10% in water and sanitation (e.g., latrines), 5% in health (including traditional midwife posts), and 2% in roads; another 26% was spent in agricultural projects including seed multiplication and communal farming; 14% in livestock or fishing (i.e., goat herding); and 17% in skills training and small business development initiatives in blacksmithing, carpentry, and soap making. Leakage of the GoBifo financial grants also appears minimal: when we asked villagers to verify detailed financial reports given to the research team by the project, community members were unable to confirm receipt for only 13.5% of the 273 transactions cross-checked.[10]

The GoBifo project is quite representative of CDD initiatives in other less developed countries. The project implementation stages—establishing a local committee, providing facilitation that aims to shift social norms, and allocating block grants—are quite standard, as is the pervasive emphasis on inclusive, transparent and participatory processes. Compared to 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 funding, GoBifo disbursed grants worth a bit under $5,000 to communities with 50 households, or 300 residents, on average (so roughly $100 per household, or $16 per capita over three and a half years). The Fearon et al. (2009) Liberia project provided roughly $20,000 to "communities" that comprised around four villages with two to three thousand residents, so $8 per capita over two years; and villages received $8,800 in Indonesia (Olken 2007). While the difference in total grant size may affect

[10]These discrepancies were of two types: i) the amounts in community records was markedly less than in project accounts; or ii) community members reported receiving building materials in kind and could not estimate their value. For each of the disputed transactions, the GoBifo accounting team produced hard copy payment vouchers signed by both a village representative (either the VDC Chair or Finance Officer) and a project field staff member.

the maximum feasible project scale, the per capita funding differences are not substantial.

3.2 Research Design

The 118 GoBifo treatment and 118 control villages were selected from a larger pool of eligible communities using a computerized random number generator. Two study districts were chosen to strike a balance in terms of regional diversity, political affiliation, and ethnic identity, while simultaneously targeting poor rural areas with limited NGO presence (see Appendix D for a map). Bombali district is located in the Northern region dominated by the Temne and Limba ethnic groups and traditionally allied with the All People's Congress (APC) political party, one of Sierra Leone's two largest parties. Bonthe district is in the South, where the Mende and Sherbro ethnic groups dominate and where the population is historically aligned with the other major party, the Sierra Leone People's Party (SLPP). Using the 2004 Population and Housing Census, the eligible pool of villages was restricted to communities considered of appropriate size for a CDD project, namely between 20 and 200 households in Bombali and 10 to 100 households in Bonthe (where villages are smaller), and once the final study sample was chosen, the villages were randomized into treatment and control groups, stratifying on ward.[11]

For each community in the study sample, government Statistics Sierra Leone staff randomly selected twelve households to be surveyed from the Census household lists. Given research interest in the dynamics of political exclusion and empowerment, the choice of respondent within each targeted household rotated among four different demographic groups in each subsequent household surveyed: non-youth male, youth male, non-youth female and youth female. All respondents are at least 18 years old, and note that the Government of Sierra Leone's definition of youth includes people up to 35 years of age (although in reality the definition of youth is a bit subjective, especially since some respondents do not know their exact

[11]We ran 500 computer randomizations and saved all resulting assignments that generated no statistically significant differences (at 95% confidence) between treatment and control groups in terms of the total number of households per village and the distance to the nearest motorable road. Among these "balanced" assignments, one was then selected at random for the final allocation of GoBifo treatment and control villages. Bruhn and McKenzie (2009) argue correctly that this process of re-randomization to achieve balance on observables may lead standard errors to be either under- or over-estimated. They show that correct inference can be achieved by including the "balancing" observables in the regression analysis as control variables, and these variables are thus included in our standard set of regression controls in all results presented below. The treatment effect estimates are thus interpreted as impacts conditional on these observables. It is worth noting, however, that coefficient estimates and standard errors are nearly unchanged whether or not these controls are included in the analysis (not shown).There were two minor data issues in measuring community size and ward location that led to a partial re-sampling of a small number of villages, however these did not affect the integrity of the randomization (see web Appendix E).

age). This data collection strategy means that for each community, and for the overall sample, responses are roughly balanced across the four demographic groups.[12]

The randomization procedure successfully generated two groups balanced along observable dimensions. Specifically, Table 1 lists the mean value in the control group and the treatment minus control pre-program difference for a variety of community characteristics (including total households, distance to nearest road, average respondent years of education, and indices for civil war exposure and local history of domestic slavery) as well as an illustrative selection of pre-program values for measures that fall under each of the three outcome "families" mentioned above. There are no statistically significant mean differences across the treatment and control groups in the 2005 values of any of these variables; Appendix F contains the same estimates for all 94 baseline measures and shows that the difference across treatment and control groups is significant at 90% confidence for only seven of these, roughly as expected by chance. Note that the analysis below typically controls for baseline values of the outcome under consideration (when it is available), addressing any incidental baseline imbalance across groups. One noteworthy pattern in the baseline data is the stark gender difference in local meeting involvement, with twice as many males (59%) than females (29%) speaking at village meetings.

3.3 Data Collection and Measurement

This analysis draws on three main data sources: household surveys from late 2005 (baseline) and mid-2009 (follow-up); village-level focus group discussions held in 2005 and 2009; and three novel structured community activities (SCAs) conducted in late 2009 shortly after GoBifo activities had ended. The SCAs were introduced with the initial post-program survey in May 2009 and then followed up in an unannounced visit five months later. The research team and enumerators were operationally separate from GoBifo staff at all stages of the project.

The 2005 household surveys collected data on baseline participation in many local collective activities, as well as detailed household demographic and socioeconomic information. To establish a panel, the field teams sought out the same respondents during the 2009 follow-up household surveys that they had previously interviewed, and the attrition rate was moderate: overall, 96% of the same households were located and 76% of the same individual respondents.

[12] These four demographic groups each comprise roughly a quarter of the adult population in these two districts in the 2004 Census (ranging from 21 to 31%), indicating that our sample is quite representative.

During the data collection visits in 2005 and 2009, the field team supervisor assembled key opinion leaders—including VDC members, the village chief, as well as women and youth leaders, among others—to describe the condition of local infrastructure and answer questions about local collective processes and activities. Research supervisors also made their own physical assessments of the quality of construction as a cross-check on focus group responses.

Given the difficulties in gauging institutional dynamics and collective action through survey responses alone, the third main type of data was gathered through the SCAs. These were designed to measure how communities respond to concrete, real-world situations requiring collective action in three different dimensions: (i) raising funds in response to a matching grant opportunity; (ii) making a community decision between two comparable alternatives; and (iii) allocating and managing an asset that was provided for free. As opposed to hypothetical vignettes or laboratory experiments in the field, these exercises more directly, realistically, and less obtrusively capture collective action outcomes of interest. SCA#1 also mimics the way outside NGOs often engage with communities, further bolstering its relevance.

SCA #1 was designed to measure whether GoBifo produced persistent effects on villages' capacity for local collective action 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 voucher was worth 50,000 Leones (roughly US$17) only if accompanied by another 100,000 Leones (US$33) from the community. Matching all six vouchers generated a sizeable 900,000 Leones, or approximately US$300, for use in the store. Take-up of the vouchers was recorded by clerks at the building materials stores. Enumerators returned to all villages five months after the initial distribution of the vouchers to assess the quality of final construction, the distribution of project contributions and benefits (i.e., did they buy zinc to build a new roof for the primary school or for the chief's home compound?), and how inclusive and transparent the management of the resulting project had been. In the context of the model, higher take up in treatment communities than in controls ($q_2 > q_0$) implies the program had a persistent effect by changing the nature of local institutions, social norms and collective action capacity (i.e., $\gamma_2 < \gamma_0$) since the financial subsidies offered through the vouchers were exactly the same for treatment and control villages.

SCA #2 was designed to measure the extent to which community decision-making is democratic and inclusive, and to assess the level of community participation. The day before

survey work, the enumerator teams met with the village head (the lowest level chiefly authority) and asked him/her to assemble the entire community for a meeting the next morning. At the subsequent meeting, the enumerators presented the community with a choice between two gifts of medium-value—a carton of batteries (useful for radios and flashlights) versus many small bags of iodized salt, each valued at roughly US$40—as a token of appreciation for participating in the research. The enumerators – who were Statistics Sierra Leone employees and not GoBifo staff – emphasized that the community itself should decide how to share the gift, and at that point the enumerators withdrew from the meeting to observe the decision-making process from the sidelines. The enumerators remained "outside" the community meeting circle and recorded how the deliberation evolved without making any comments of their own. Among other things, the enumerators recorded who participated in any side-meetings; the degree to which the chief, village head and elders dominated the discussion; the extent of debate in terms of time and the number of comments; and a subjective assessment of the apparent influence of different sub-groups (e.g., women) on the final outcome. This exercise thus provides concrete quantitative data on the relative frequency of female versus male speakers, and youth versus non-youth speakers in an actual community meeting.[13] Note that these are exactly the same metrics that the GoBifo facilitators were required to track during project meetings as part of their own internal performance assessment, where the facilitators reported real strides in women's and youth participation, leadership and power in treatment communities (GoBifo Project 2008). Returning to the model, a durable program effect on the benefits of civic engagement would imply that women and youths were more active participants in the treatment communities ($\omega_2 > \omega_0$).

SCA #3 was designed to gauge the extent of elite capture of resources as a result of the CDD process, as well as the nature of broader collective action processes. During the same follow-up visit in 2009, the enumerators also gave each village a large plastic tarpaulin sheet as a gift. Tarpaulins are frequently used in Sierra Leone as makeshift building materials for roofing, and in agriculture as a surface for drying grains or as a covering to protect them from rain. During the second 2009 follow-up visit five months later, enumerators recorded which households had had access to the tarpaulin in the intervening period. This activity also captures an element of collective action, as enumerators assessed whether villages had been able to decide

[13] Of the four enumerators, one focused their data collection on the participation of youths, one on women, one on all adults and the fourth kept careful track of each person who spoke publicly.

on a use for the tarp at all, and whether it had been put mainly towards a public end (e.g., use as a communal grain drying floor) or private ends (patching the roof of an individual's home).

3.4 Econometric Specifications

In what follows, we present results for the specific hypotheses described in our *ex ante* analysis plan, a document that was finalized before we analyzed any follow-up data. The genesis of the plan was a pre-program 2005 agreement between the research and project teams that set out the areas GoBifo was likely to impact 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 eleven hypothesized areas of impact and the econometric specifications to be used, which we archived with the Abdul Latif Jameel Poverty Action Lab randomized evaluation archive in August 2009. This approach limits data mining, or an *ex post* rationalization 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 econometric models (Leamer 1974) before data analysis began. While the randomized 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 jointly identify the different hypotheses laid out in the 2005 document (see O'Brien 1984; Kling and Liebman 2004; Anderson 2008). 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 supplementary web Appendix G.

There are two minor deviations from the original *ex ante* analysis plan in what we present below. We added a twelfth hypothesis (called hypothesis 1 below) by pulling together project implementation outcomes that had already been explicitly included as outcomes within the

original eleven hypotheses. Thus no new outcome measures were added or excluded in what we present below. Those who wish to consider only the results as exactly laid out *ex ante* can ignore hypothesis 1. However, we feel it was an oversight to exclude a project implementation hypothesis beforehand and thus still find the results of hypothesis 1 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:

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

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 randomizations; W_c is a fixed effect for geographic ward, the administrative level on which the randomization was stratified; and ε_c is the usual idiosyncratic error term. Elements of X_c include distance from road, 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 β_1, the average treatment effect. Note that while some outcomes are measured at the household (e.g., radio ownership) or individual level (e.g., political attitudes), for simplicity we measure all 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 in 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 β_1, the average treatment effect, and here the disturbance terms are clustered at the village level. Results are robust to the exclusion of the vector of community controls and to limiting our analysis to only the post-program data (as shown in the sparse specifications in supplementary web Appendix G). 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 set out in our analysis plan). As we do not find any evidence for heterogeneous effects along any of these dimensions, we have excluded this discussion from the main text (see Appendix H for the results by outcome family).

The mean effects index for a hypothesis captures the average relationship between the GoBifo treatment and the K different outcome measures grouped in that hypothesis. Following Kling and Liebman (2004), estimation of the index first standardizes outcome variables into comparable units by translating each one into standard deviation units (by subtracting the mean and dividing by the standard error of the control group) before regressing each outcome on the vector of independent variables. The index coefficient is the mean of these K standardized treatment effects. The estimation method calculates the standard error of the index itself, which depends on both the variances of each individual $\beta_{1,k}$ as well as any covariances between $\beta_{1,k}$ and $\beta_{1,\neg k}$, requiring a seemingly unrelated regressions (SUR) system approach to test the cross-equation hypothesis that the average index of K coefficients equals zero.

4. Empirical Results

Table 2 presents a concise summary of the mean effect results for all twelve hypotheses, grouped into the three outcome families (see web Appendix I for a graphical analogue). The positive and significant (at 99% confidence) mean effect estimate for family A (hypothesis 1) indicates that GoBifo achieved its most immediate objective of creating the organizations and linkages that could facilitate local collective action. Specifically, the coefficient indicates that GoBifo increased these measures by 0.687 standard deviations on average. This indicates that GoBifo was quite well-implemented, perhaps more so than many other real-world development projects.

Given the strong implementation performance, the most immediate impacts one might expect are on measures of local infrastructure and other hardware. The large positive estimated mean effect for family B implies that GoBifo led to a 0.273 standard deviation unit average increase in such measures. This reflects positive effects on the stock and quality of local public goods, where we find a significant increase of 0.164 (hypothesis 2); and on measures of general economic welfare, at 0.399 (hypothesis 3). Reflecting back on the theoretical framework, these increases provide strong support for the prediction that the combination of lowering the marginal

cost of public goods through grants, as well as reducing coordination costs through the establishment of new institutions, led to greater public investment during project implementation (in the notation of the model, $q_1 > q_0$). The next question is how much of this effect was driven by changes in institutions, norms and collective action capacity.

The small and not statistically significant mean effect estimate for family C suggests that the experience of working together in GoBifo, and the introduction of new institutions and processes, did not durably change the nature of local collective action. The program's democratic decision-making and "help yourself" approach did not appear to spill over into other realms of village life nor to persist into the post-program period. We find no evidence that GoBifo led to fundamental changes in local capacity to raise funds and act collectively outside of the project, the nature of decision-making, the influence of women or youths, or a range of social capital outcomes. In the context of the model, these null results suggest that GoBifo did not permanently increase the benefits of civic engagement for marginalized groups (i.e., $\omega_2 = \omega_0$), and that the organizing institutions established did not persistently reduce the fixed costs of collective action ($\gamma_2 = \gamma_0$). In the subsections that follow, we flesh out these results with an illustrative sample of outcomes under each family; those interested in any particular outcome omitted from the discussion below (due to space constraints) should refer to Appendix G for the entire inventory of results for all twelve hypotheses.

4.1 Family A: Project Implementation and Direct Program Goals

The first family measures the extent to which the project established Village Development Committees (VDCs); helped communities draw up development plans and open bank accounts; and created links between the villages and their local government representatives. The first two panels of Table 3 present results for several outcomes in this family, where the first seven "full sample" outcomes in Panel A apply to all communities within the sample; while the remaining six "conditional" outcomes in Panel B are conditioned on the existence of public infrastructure and thus only apply to those communities that have the particular good. All of these treatment effects are greater than zero and nine are statistically significant at 95% confidence.

Regarding interpretation, the treatment effect estimate from the panel specification in the first row of Table3indicates an increase of 34.1 percentage points in the existence of a VDC. VDCs already existed in many Sierra Leonean villages when GoBifo was launched, having been

introduced by humanitarian assistance groups during the war-torn 1990's (Richards et al. 2004). By the post-program period, 86.3% of GoBifo communities had a VDC compared to 45.8% of controls, a large effect. The corresponding coefficient in the second row indicates that GoBifo increased the likelihood that a community was visited by a member of its Ward Development Committee in the past year by 15.6 percentage points. Row 3 shows a positive treatment effect on the existence of village development plans by 29.6 percentage points, nearly a 50% increase on the base of 61.7% in the controls. Row 4 reveals an increase in having a village bank account of 70.6 percentage points, capturing a tenfold increase. In Panel B, the household survey asked respondents whether a member of the Ward Development Committee or district council was "directly involved in the planning, construction, maintenance or oversight" of several local public goods. Note that the treatment effect is positive and significant for nearly all outcomes. This suggests that GoBifo successfully led local politicians to increase their involvement in village projects, consistent with its objective of supporting the broader decentralization process.

4.2 Family B: Impacts on "Hardware": Local Public Goods and Economic Activity

Outcome family B explores the impact of project activities on public goods and economic welfare. As communities used the bulk of GoBifo grants on local infrastructure, hypothesis 2 explores treatment effects on the quantity and quality of local public goods. In addition, roughly one sixth of the grants were used to launch projects dedicated to job skills training or small business development—such as carpentry, soap-making and seed multiplication initiatives—that, if well implemented, could translate into higher small business profits, and perhaps lead to sustainably higher future earnings. Moreover, GoBifo injected cash grants into very poor communities, and as with any assistance, a portion of the funds are surely fungible. Via potentially all of these mechanisms, hypothesis 3 considers project impacts on measures of community-wide economic activity and household welfare.

While combining measures within a single hypothesis into sub-indices was not specified in our *ex ante* analysis plan, the outcomes under hypothesis 2 naturally form three sub-groups: the stock of local public goods, the quality of such goods, and community financial contributions to their construction and upkeep. Regarding the stock, the first five rows of Panel C in Table 3 present impacts for an illustrative sample of goods. Note that four of these treatment effects are positive and three are statistically significant. Specifically, there are marked increases in the

proportion of villages with a functional traditional midwife post by 17.5 percentage points, community center by 24.1, and latrine by 21.0. Calculating a mean index on the entire sub-group reveals a highly significant increase of 0.258 standard deviations (s.e.0.049, not shown).[14]

Turning to the next sub-group, the first three rows of Panel D show positive GoBifo impacts on the quality of construction of three of the most common public goods—primary schools, latrines and grain drying floors. The effects are all positive as is the quality sub-group index overall, which shows an increase of 0.296 standard deviation units (s.e. 0.077).[15] These measures combine impacts from the GoBifo funded infrastructure projects, as well as any potential effects from better local collective action in maintaining existing infrastructure. However, as there is no evidence that management practices did in fact change in treatment villages, the leading interpretation is that the positive impacts are being driven by the grants.

The last three rows of Panel D present illustrative results for the final sub-group of outcomes that concern community financial contributions to existing infrastructure. Two of these are negative and one is statistically significant. Looking across nine different local public goods, the sub-group index is negative (at -0.113 standard deviations) but not statistically significant (s.e. 0.104). Combined with the negative and marginally significant treatment effect on whether the community approached another NGO or donor for financial support (in row 15 of Table 3), these provide suggestive evidence that GoBifo funds served as a substitute, rather than a complement, for the community's own resources. At a minimum, they indicate that the GoBifo grants did not serve as a catalyst for additional external fund-raising nor did project experiences encourage participants to seek out further development assistance beyond the project itself. The SCA findings (discussed in Section 4.3 below) reinforce this finding.

The second hypothesis in the "hardware" category concerns measures of general economic welfare at the household and community levels. The first two outcomes in Panel E of Table 3 refer to village-level outcomes, where we see a 30% increase in the number of petty traders in row 25 (0.7 more traders on a base of 2.4 traders in the control group) and a 13% increase in goods locally available for sale in row 26. The last four outcomes are aggregated from household survey reports. We observe improvements in an asset ownership score (derived using principal components analysis) in row 27, where the underlying assets include common

[14] The mean effect index for the sub-group includes impacts on six additional goods not p resented due to space constraints: water wells, peripheral health unit, market, grain store, sports field and sports uniforms.
[15] The quality of construction sub-group index uses two measures for each good and includes effects on water wells.

household durables (e.g., 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% in control communities (row 29). We find no impact on 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 30).

In sum, the positive treatment effects for outcome family B suggest that investment in local public goods did increase substantially during the project as predicted by the theoretical model. To determine the role played by more effective local institutions (versus the block grants), we next examine post-program outcomes after the block grants had been spent.

4.3 Family C: Impacts on "Software": Local Institutions and Norms for Collective Action

The first hypothesis under the software family (hypothesis 4) covers outcomes relating to collective action and contributions to local public goods. The mean effect for this hypothesis is not statistically distinguishable from zero (0.041 standard deviations with a standard error of 0.042); and of the 59 full sample and conditional outcomes evaluated, only seven treatment effects are significant at 95% confidence, with five positive in sign and two negative. The subset of outcomes relating to the matching grant opportunity (SCA#1) provides the most succinct and concrete illustration of the lack of program impacts in this area. The top panel of Table 4 shows that there was no differential take-up of the subsidized building vouchers: 62 treatment (52.5%) and 64 control villages (54.2%) redeemed vouchers at local supply stores; nor is there any difference in the number of vouchers redeemed, as most of the villages that cashed in any vouchers used all six. The ability to mobilize around a new opportunity and the willingness to raise funds for local infrastructure is close to the essence of local collective action in the model. This finding implies that the program did not have durable effects on collective action capacity.

Other outcomes under this hypothesis consider household contributions to existing local public goods, where we expand the set of contributions to include labor, local materials, or food for project workers, yet continue to find no treatment effect. There are also no differences in contributions to several different local self-help groups (i.e., rotating savings groups and labor gangs) nor in financial support of community teachers. Lastly, while treatment villages were more likely to have a communal farm, by 23 percentage points (significant at 99% confidence),

the total number of respondents in treatment areas who had worked on a communal farm in the past year was no higher. This presents a telling example of how project-funded activities—for example, the subsidized provision of seeds and tools for a community farm—exerted a proximate effect on the establishment of a local organization established to capture that funding, but did not have any lasting impacts on actual communal cultivation in subsequent years.

These findings raise troubling questions about GoBifo's long term impacts. Clearly, community members gained experience in working together to successfully implement local development projects over the nearly four years of the project. Yet their GoBifo-specific experiences did not lead to greater capacity to take advantage of new opportunities that arose after the program ended. Most strikingly, while GoBifo often created new structures designed to facilitate local development—the VDC, a development plan, a bank account, and a communal farm—all designed to reduce local collective costs (γ_t in the theoretical framework) the program left them no better able to take advantage of the realistic matching grant opportunity in SCA #1.

The second "software" hypothesis includes outcomes relating to the civic involvement of socially marginalized groups. Since the inclusion of women and youth held great prominence in GoBifo's objectives and facilitator operating manuals, it also received special attention in the data collection. Covering an exhaustive battery of measures, the mean effect is a precisely estimated zero (see hypothesis 5 in Table 2) indicating no overall impact on the role of women or youth in local decision-making, or on the transparency and accountability of decision-making more generally. Of the 72 distinct outcomes considered, only six were statistically significant at 95% confidence, dividing equally between positive and negative treatment effects.

Enumerator observations during SCA #2, when villages met to decide between salt versus batteries, provide a clear illustration of this zero result. In Panel B of Table 4 there are 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 (s.e. 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 baseline. We similarly find no impact on 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-process appeared to the enumerators (e.g., by holding a direct vote). These results are further substantiated by respondent reports recorded immediately after the meeting of how the tarpaulin allocation choice in SCA #3 was made, including which individuals had the final "say" and to what extent the decision was dominated by local elites (i.e., village headmen and male elders). Moreover, respondent opinions collected during the second 2009 follow-up survey (five months later) also find no treatment effects on reports about how decisions were made to distribute the salt or batteries (SCA #2); how to use the tarp (SCA #3); whether to raise funds for the building materials vouchers, and if so, how to mobilize funds, which items to purchase, and how to manage any construction (SCA #1).

Despite all of the effort in GoBifo to elevate the position of women and youth, we thus do not observe any improvement in their role relative to older men in community decision making. Even for relatively low cost actions like speaking up in meetings, the nearly four years of GoBifo project activities did not translate into greater apparent voice for marginalized groups. In the context of the theory, this suggests no persistent gains in the individual benefits of participation for these groups (b_{it} in the model), and provides additional evidence that the increase in public investment observed during project implementation was likely driven by the financial subsidy rather than fundamental changes in local institutions or *de facto* power.[16]

The third software hypothesis (which was not an explicit aim of the project) asks whether by espousing more democratic ways of managing local development, the project led to changes in the role of the traditional chiefly authorities. Taking all outcomes together, the mean effect for hypothesis 6 is also zero (Table 2). Most outcomes under this hypothesis estimate the extent to which the village head and elders dominated the SCA decisions. While we find large variation in how these decisions are made—at one extreme, in two villages the Chief decided between the salt and 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—as mentioned above, we find no systematic differences in averages across treatment and control villages.

A leading explanation (with some support in the data) is that elites exerted substantial control over the new organizations GoBifo created. As an example, the composition of VDC members and leaders is no different in treatment and control villages (for the roughly half of

[16] We cannot rule out that the subsidy was particularly effective (i.e., led to such notable increases in public goods) in part because of the project's facilitation and emphasis on participation and transparency.

control communities with a VDC in 2009). Specifically, in both only 21% of VDC members are women and only 26% are 35 years old or younger, while 88% of VDC chairs are men and 87% are older than 35, and traditional chiefdom authorities and elders account for 52% of the chairs; the only notable difference is that women are more likely to hold the treasurer position in GoBifo villages: 57% in treatment versus only 31% in control areas. These patterns highlight a tension inherent in the CDD approach: leveraging the capacity of existing institutions may be expedient for immediate project implementation while simultaneously limiting the likelihood of fundamental institutional transformation or changes in *de facto* power for marginalized groups.

We therefore tested the related hypothesis that CDD may enable local elites to capture a disproportionate share of economic benefits. We explored this issue directly by measuring the allocation of a new public asset—the tarp (SCA #3)—in villages during the first 2009 follow-up visit, and then observing how it was being used in the unannounced second 2009 visit five months later. While the analysis finds no treatment effects on the extent of elite capture, it also reveals that the level of elite capture is, perhaps surprisingly, relatively low in the study communities. Panel C of Table 4 shows that for the 90% of communities that had used the tarp by the time of the second visit, 86% had put the tarp towards a public purpose, such as a communal rice drying floor or local ceremony. The most obvious example of elite capture would be use of the tarp to patch the roof of a single individual's house, which happened in fewer than 3% of all villages. That said, several communities had not yet used the tarp and were storing it at a private residence, which either suggests a failure to agree on the appropriate way forward, or signals the risk of future elite capture, or both.

The next three hypotheses explore proxies for "social capital"—self-expressed trust of others (hypothesis 7), involvement in local groups and networks (hypothesis 8), and access to information (hypothesis 9)—emphasized alongside collective action and inclusion in the official GoBifo project objectives (World Bank 2004, GoBifo 2007). Despite exploring a wealth of measures, the analysis reveals no treatment effects on social capital and the three mean effects indices are all indistinguishable from zero. Beginning with trust, the only significant effect is an increase in reported trust of NGOs and donor projects: residents in treatment communities were 5.4 percentage points more likely to agree that NGOs or donors "can be believed" (close to the Krio translation for trust) as opposed to you "have to be careful" in dealing with them. There are no effects on the remaining eleven indicators, which combine respondent self-reports regarding

how much they trust various groups with concrete examples of trusting behavior, such as entrusting money to a neighbor to purchase market goods on your behalf.

Second, enumerators asked respondents whether they were a member of a local self-help group (such as a credit/savings group, communal labor gang, school committee, funeral savings group, fishing cooperative, women's group or youth group, among others) and if so, whether they had attended a meeting and contributed financially or in labor in the past month (hypothesis 8). We find no significant treatment effects on these indicators nor on other measures of local cooperation, such as whether the respondent had helped a neighbor re-thatch the roof of their house, a time-intensive activity that one cannot easily do alone.

There is also no evidence of treatment effects on households' access to information about local government or governance (hypothesis 9). Among 21 outcomes, only one—the proportion of villages visited by a WDC member, discussed above—shows statistically significant effects. The collection of zero effects includes measures of how much respondents know about what the community is doing with the building vouchers (SCA #1) and tarp (SCA #3); whether they can name their district council and chiefdom leaders; and their ability to answer objective questions about how local taxes are collected and used.

While the mean effect index for participation in local governance in Table 2 (hypothesis 10) is positive and statistically significant, it is largely driven by the outcomes already discussed under family A. Specifically, we find large impacts on the existence of VDCs and village plans, and increases in the oversight of local public goods by chiefdom authorities that mirror earlier results on the involvement of local government representatives. There is no evidence, however, that these stronger links with either set of local officials translated into more active individual political engagement, such as self-reported voting or running for local office. Similarly, treatment communities were no more likely to use the building materials and tarp in the SCAs for goals specified in their village development plan. Reinforcing earlier results, this disconnect between the articulation of a development plan and its real-world application suggests that few communities applied GoBifo project tools to initiatives beyond the program.

There are no impacts on crime and conflict in treatment villages or in the mechanisms through which they are resolved, leading to a zero mean effect for hypothesis 11 (Table 2). Of the ten indicators considered, only one—the 2 percentage point reduction in household reports of physical fighting over the past year—is significant at 95% confidence. While the nine null

results imply that project efforts to enhance conflict management capacity may not have created lingering benefits, on the positive side it provides some reassurance that the infusion of block grants into the treatment communities at least did not spark increased conflict.

The twelfth and final hypothesis concerns the nature of individual political and social attitudes. The GoBifo program's emphasis on the empowerment of women and youth, and the transparency of local institutions, may have engendered a more equitable or "progressive" outlook toward politics and society more generally. Even if there are no changes in actual decision-making processes or local collective outcomes (as above), a marked change in expressed attitudes might still mean that the seeds for future social change have been planted by the program. Enumerators gauged attitudes using pairs of opposing statements, such as "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 asking respondents which they agreed with more. These paired statements capture respondent views on a diverse range of topics including the acceptability of the use of violence in politics (a particularly salient issue in post-war Sierra Leone), of domestic violence, of youth and women in leadership roles, paying bribes, and coerced labor. Once again, there are no significant program effects, despite the concern that social desirability bias might lead some respondents to express views promoted by the program. The only significant impact is a positive 3.8 percentage point increase in agreement with the statement that young people can be good leaders. However, recall that this change in opinions did not translate into more youths holding actual leadership positions on the VDC by 2009, or to more youth participation in the SCA meetings. Attitudinal change may be a necessary step toward changing future behavior, but almost four years of an intensive community driven development program did not lead to detectable changes in a wide array of expressed attitudes.

4.4 Robustness and Validity Checks

This section evaluates the robustness of the results. To start, we consider typical threats to randomized experiments. Fortunately, there were no problems with treatment non-compliance: all communities assigned to the treatment group received the program and none of those in the control group participated; and respondent attrition rates are no different in treatment and control areas. The baseline statistics presented in Table 1 and supplementary Appendix F also suggest that the randomization process successfully created two groups of villages that were similar

along a wide range of observables. Note further that the analyses use the baseline value of the outcome of interest as a control variable wherever such panel data is available. Thus in order for spurious differences between the two groups to explain the positive impacts, the treatment group would on average have had to be on a different trajectory than the controls, but there is no reason to believe this should systematically be the case given the randomized research design.

We next consider reasons why the treatment effect estimates might be underestimates. First, significant program spillovers from treatment to control communities could lead us to underestimate program impacts, since the control communities would also be receiving program benefits, albeit indirectly. For this to be true, we might expect the coefficient on the "post" indicator ($POST_t$ in equation 12) to often be positive and significant, but this is not the case: across all the outcomes in Appendix G where panel data is available, there are exactly as many (21) positive as negative coefficient estimates on $POST_t$ that are statistically significant, and thus it seems unlikely that the results are biased by spillovers across communities.

A further concern is that the projects GoBifo simultaneously implemented at the ward level systematically benefited the control group at the expense of the treatment group. There was a separate pot of funding for each ward that was allocated by the Ward Development Committee (see footnote 1). Bias could result if WDC members took into account the placement of GoBifo village-level projects in deciding where to locate the ward projects and targeted those areas that had not already benefited, perhaps as a way of compensating them for losing out on village-level assistance. However, there are no meaningful differences in the targeting of ward-level projects across treatment and control villages, and, if anything, treatment villages are slightly more likely to benefit: while 15.2% of respondents in treatment areas reported that a household member benefited from a ward-level project, only 6.1% of respondents in control areas reported benefits.

A final concern is that the outcome measures were simply insufficiently refined to detect subtle decision-making, institutional, political or social differences between treatment and control communities. While some of our measures are certainly better than others, our main strength lies in the diversity and multiplicity of measures we use and the fact that they all produce similar results. As an example, we combine different data collection approaches, employing both survey self-reports on the percentage of female and male respondents who spoke during the SCA meetings with direct enumerator observation of how many men and women they saw speaking during the meeting. The research teams also gathered information from a variety

of sources, interviewing men and women in their own homes, holding focus group discussions with key opinion leaders, observing a community decision as it unfolded, and recording their own independent assessment of the construction quality of local infrastructure. Lastly, we examine a large number of outcomes. Taking all these data together, the "zero" GoBifo program effects are quite precisely estimated. To illustrate, the maximum true positive treatment effect on the proportion of women speaking (in the salt versus battery SCA #2 deliberation) that we may have incorrectly ruled out at 95% confidence is one additional female speaker per every 4.3 villages we visited, which is quite small. In the mean effects analysis, which combines many outcome measures, confidence intervals are considerably tighter.

5. Conclusion

A well-implemented community driven development (CDD) program in Sierra Leone was successful at setting up new village structures and improving the stock of local public goods, but did not lead to any lasting changes in village institutions, local collective action capacity, social norms and attitudes, or the nature of *de facto* political power. A large number of collective action outcomes recorded at the end of the program and in the immediate post-program period tell a consistent story that participation in the CDD project did not catalyze greater collective action or voluntary contributions to public goods in a sustainable fashion.

The results contradict the currently popular notion in foreign aid circles that CDD is an effective method to initiate social change or fundamentally alter local decision-making processes. Despite the project's intensive community organizing component focused on strengthening the role of women and young men, nearly four years later we see that women and youths are no more likely to voice opinions about how the community should manage new public assets. Exposure to a CDD process similarly did not make traditional elites more willing to seek out the views of others in making community decisions, nor were villages any better able to raise funds in response to a matching grant opportunity. While "good" institutions may be critical for successful economic development, our findings provide another piece of evidence that institutions and social norms are difficult to change.

Our results also challenge the aid pessimist's view that external assistance cannot improve the lives of the poor in countries with weak institutions. While we should not be so naïve as to think that structural factors like social organization and institutions are easily

transformed (Easterly 2006, Kremer and Miguel 2007), we find that well allocated external aid can have a positive impact on welfare (Sachs 2005). Indeed our results suggest that the comparative advantage of the World Bank and other donors may lie more in providing development "hardware," and less in instigating large-scale institutional and social change, at least not using current tools such as CDD. Importantly, however, we cannot rule out that part of GoBifo's success in delivering hardware impacts was due to its emphasis on transparency and the inclusion of marginalized groups. Returning to the comparison between informal interventions focused on reshaping norms, like CDD, and changes to the rules of formal institutions, like female leadership quotas, the limited existing evidence suggests that the latter may be a more effective way to alter *de facto* power dynamics and social perceptions in a modest timeframe (Duflo and Chattopadhyay 2004; Beaman et al. 2009).

As our results concern one program in one country, these more general implications are clearly speculative. However, we can conclude with certainty that far more research is needed to identify the precise reforms and external interventions that can successfully reshape institutions to enhance collective action capacity while promoting accountability and inclusion.

REFERENCES

Acemoglu, Daron, Simon Johnson and James A. Robinson. 2001. "The Colonial Origins of Comparative Development: An Empirical Investigation." *American Economic Review*, 91(5): 1369-1401.

Acemoglu, Daron and James A. Robinson. 2008. "Persistence of Power, Elites, and Institutions." *American Economic Review,* 98(1): 267-293.

Alesina, Alberto, Reza Baqir and William Easterly. 1999. "Public Goods and Ethnic Divisions." *Quarterly Journal of Economics*, 114: 1243-1284.

Anderson, Michael L. 2008. "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedaian, Perry Preschool, and Early Training Projects." *Journal of the American Statistical Association,* 103(484):1481-1495.

Banerjee, Abhijit V., Rukmini Banerji, Esther Duflo, Rachel Glennerster & Stuti Khemani. 2010. "Pitfalls of Participatory Programs: Evidence from a Randomized Evaluation in Education in India." *American Economic Journal: Economic Policy*, 2(1): 1-30.

Banerjee, Abhijit and Lakshmi Iyer. 2005. "History, Institutions, and Economic Performance: The Legacy of Colonial Land Tenure Systems in India." *American Economic Review,* 95(4): 1190-1213.

Bardhan, Pranab. 2002. "Decentralization of Governance and Development." *The Journal of Economic Perspectives*, 16(4): 185-205.

Beaman, Lori, Raghabendra Chattopadhyay, Esther Duflo, Rohini Pande, and Petia Topalova. 2009. "Powerful Women: Does Exposure Reduce Bias?" *Quarterly Journal of*

Economics, 124(4): 1497-1540.
Bjorkman, Martina and Jakob Svensson. 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.
Bruhn, Miriam, and David McKenzie. 2009. "In pursuit of balance: Randomization in practice in development field experiments."*American Economic Journal: Applied Economics*, 1(4): 200-232.
Dal Bó, Pedro, Andrew Foster and Louis Putterman. 2010. "Institutions and Behavior: Experimental Evidence on the Effects of Democracy." *American Economic Review*, 100(5): 2205-2229.
Dongier, Philippe, Julie Van Domelen, Elinor Ostrom, Andrea Rizvi, Wendy Wakeman, Anthony Bebbington, Sabina Alkire, Talib Esmail and Margaret Polski.2003. "Chapter 9: Community-Driven Development." In *The Poverty Reduction Strategy Sourcebook Volume 1*, 301-331.Washington DC: The World Bank.
Duflo, Esther and Raghabendra Chattopadhyay. 2004. "Women as Policy Makers: Evidence from a Randomized Policy Experiment in India." *Econometrica,* 72(5): 1409-1443.
Easterly, William. 2006. *The White Man's Burden: Why the West's Efforts to Aid the Rest Have Done So Little.* New York: Penguin.
Engerman, Stanley L. and Kenneth L. Sokolof. 1997. "Factor Endowments, Institutions, and Differential Paths of Growth Among New World Economies: A View from Economic Historians of the United States." In *How Latin America Fell Behind*. Stanford: Stanford University Press.
Fearon, James, Macartan Humphreys and Jeremy M. Weinstein. 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.
Glennerster, Rachel, Edward Miguel and Alexander Rothenberg. 2009. "Collective Action in Diverse Sierra Leone Communities." University of California, Berkeley Working Paper.
GoBifo Project. 2006. "GoBifo Overall Budget." Project mimeograph.
GoBifo Project. 2007. "Operations Manual: Version 2 June 2007." Project mimeograph.
GoBifo Project. 2008. "Results towards Goals and Objectives." Project mimeograph.
Gugerty, Mary Kay and Michael Kremer. 2008. "Outside Funding and the Dynamics of Participation in Community Associations." *American Journal of Political Science,* 52(3):585-602.
Horton, R. and R. Smith. 1999. "Time to register randomized trials." *British Medical Journal*, 319:865.
Khwaja, Asim I. 2004. "Is Increasing Community Participation Always a Good Thing?" *Journal of the European Economic Association,* 2(2/3): 427-436.
Kling, Jeffrey R. and Jeffrey B. Liebman. 2004. "Experimental Analysis of Neighborhood Effects on Youth." Princeton University Manuscript.
Keen, David. 2003. "Greedy Elites, Dwindling Resources, Alienated Youths: The Anatomy of Protracted Violence in Sierra Leone." *International Politics and Society,* 2: 67-94.
Kremer, Michael and Edward Miguel. 2007."The Illusion of Sustainability." *Quarterly Journal of Economics*, 122(3): 1007-1065.
Labonne, Julien and Robert Chase. 2008. "Do Community-Driven Development Projects Enhance Social Capital? Evidence from the Philippines."The World Bank Policy Research Working Paper 4678.

Leamer, Edward E. 1974. "False Models and Post-Data Model Construction." *Journal of the American Statistical Association*, 69(345):122-131.
Leamer, Edward E. 1983. "Let's Take the Con Out of Econometrics." *American Economic Review,* 73(1): 31-43.
Mamdani, Mahmood. 1996.*Citizen and Subject: Contemporary Africa and the Legacy of Late Colonialism*. Princeton, NJ: Princeton University Press.
Mansuri, Ghazala and VijaynedraRao. 2004. "Community-Based and –Driven Development: A Critical Review." *The World Bank Research Observer,* 19(1): 1-39.
Oates, Wallace E. 1999. "An Essay on Fiscal Federalism."*Journal of Economic Literature*, 37(3):1120-1149.
O'Brien, Peter C. 1984. "Procedures for Comparing Samples with Multiple Endpoints." *Biometrics,* 40(4):1079-1087.
Olken, Benjamin A. 2007."Monitoring Corruption: Evidence from a Field Experiment in Indonesia."*Journal of Political Economy,* 115(2): 200-249.
Olken, Benjamin A. 2010. "Direct Democracy and Local Public Goods: Evidence from a Field Experiment in Indonesia." *American Political Science Review*, 104(2):243-267.
Olson, Mancur. 1982.*The Rise and Decline of Nations.* New Haven: Yale University Press.
Osborne, Martin J., Jeffrey S. Rosenthal and Matthew A. Turner. 2000. "Meetings with Costly Participation." *American Economic Review*, 90(4): 927-943.
Reno, William. 1995., *Corruption and State Politics in Sierra Leone.* Cambridge and New York: Cambridge University Press.
Richards, Paul. 1996. *Fighting for the Rainforest: War, Youth and Resources in Sierra Leone*. London: James Currey & Portsmouth, NH: Heinemann for the International African Institute.
Richards, Paul, Khadija Bah and James Vincent. 2004. "The Social Assessment Study: Community-driven Development and Social Capital in Post-war Sierra Leone." Unpublished manuscript.
Sachs, Jeffrey D. 2005.*The End of Poverty: Economic Possibilities for Our Time.* New York: Penguin Press.
Simes, R.J. 1986. "Publication bias: The case for an international registry of clinical trials", *Journal of Clinical Oncology*, 4:1529-1541.
Voss, John.2008. "Impact Evaluation of the Second Phase of the Kecamatan Development Program in Indonesia." The World Bank, Jakarta Research Paper.
United Nations. 2004.*Human Development Report 2004*. New York: United Nations Development Program, Oxford University Press.
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, DC: The World Bank.

Table 1: Baseline (2005) Comparison between Treatment and Control Communities

	Baseline mean for controls	T-C difference at baseline	N
	(1)	(2)	(3)
Panel A: Community Characteristics			
Total households per community	46.76	0.30	236
		(3.67)	
Distance to nearest motorable road in miles	2.99	-0.32	236
		(0.36)	
Index of war exposure (range 0 to 1)	0.68	-0.01	236
		(0.02)	
Historical legacy of domestic slavery (range 0 to 1)	0.36	0.03	236
		(0.06)	
Average respondent years of education	1.65	0.11	235
		(0.13)	
Panel B: Selected Outcomes from Project Implementation Family A			
Proportion of communities with a Village development committee (VDC)	0.55	0.06	232
		(0.06)	
Proportion visited by Ward Development Committee (WDC) member in past year	0.15	-0.01	228
		(0.05)	
Panel C: Selected Outcomes from "Hardware" Family B			
Proportion of communities with a functional grain drying floor	0.23	0.05	231
		(0.05)	
Proportion of communities with a functional primary school	0.41	0.08	230
		(0.06)	
Average household asset score	-0.06	0.11	235
		(0.08)	
Supervisor assessment that community is "better off" than others nearby	0.31	0.04	201
		(0.06)	
Proportion of communities with any petty traders	0.54	-0.01	226
		(0.06)	
Panel D: Selected Outcomes from "Software" Family C			
Respondent agrees that chiefdom officials can be trusted	0.66	-0.01	235
		(0.02)	
Respondent agrees that Local Councillors can be trusted	0.61	0.00	235
		(0.02)	
Respondent is a member of credit / savings group	0.25	-0.03	235
		(0.02)	
Respondent is a member of labor sharing gang	0.50	-0.01	235
		(0.03)	
Among males who attended a community meeting, respondent spoke publicly	0.59	-0.02	235
		(0.04)	
Among females who attended a community meeting, respondent spoke publicly	0.29	0.03	229
		(0.04)	
Respondent claimed to have voted in last local elections	0.85	-0.01	235
		(0.02)	

Notes on table: i) significance levels indicated by + p <0.10, * p <0.05, ** p <0.01; ii) robust standard errors; iii) the T-C difference is the pre-program "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 randomization (total households and distance to road); iv) regressions for the two balancing variables in rows 1 and 2 exclude the outcome from the set of controls; and v) see Appendix F for the T-C difference for all 94 outcomes collected in the baseline survey.

Table 2: Summary of GoBifo Program Impacts by Research Hypothesis and Outcome Family

Hypotheses by family	GoBifo Mean Effect (std. error)
Family A: Project Implementation	
Mean Effect for Family A (Hypothesis 1; 7 total outcomes)	**0.687**** **(0.062)**
H1: GoBifo creates functional development committees (7 outcomes)	0.687** (0.062)
Family B: Development Infrastructure Impacts or "Hardware" Effects	
Mean Effect for Family B (Hypotheses 2 and 3; 30 total outcomes)	**0.273**** **(0.032)**
H2: GoBifo increases the quality and quantity of 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	
Mean Effect for Family C (Hypotheses 4, 5, 6, 7, 8, 9, 10, 11 and 12; 146 total outcomes)	**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)

Notes on table: i) significance levels indicated by + p <0.10, * p <0.05, ** p <0.01; ii) robust standard errors clustered by village for panel data; iii) includes fixed effects for the district council wards (the unit of stratification); iv) 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; v) these mean effect estimates are limited to the full sample set of outcomes that excludes all conditional outcomes (i.e. those that depend on the state of another variable--for example, quality of infrastructure depends on the existence of the infrastructure); and vi) for the complete list of all full sample and conditional variables under each hypothesis--including the exact wording of survey questions and treatment effect estimates for each distinct outcome measure--see Appendix G.

Table 3: Illustrative Treatment Effect Estimates for Selected Outcome Measures

Row	Outcome variable		Mean in Controls	Treatment Effect	Standard Error	N	Specification
	Panel A: Hypothesis 1 - Full Sample Outcomes						
1	Village development committee		0.458	0.341**	(0.077)	467	Panel
2	Visit by WDC member		0.212	0.156*	(0.070)	462	Panel
3	Village development plan		0.617	0.296**	(0.048)	221	Cross section
4	Community bank account		0.081	0.706**	(0.045)	226	Cross section
	Panel B: Hypothesis 1 - Conditional Outcomes						
5	[Given functional infrastructure in the village] A WDC or district 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
6		Grain drying floor	0.243	0.140*	(0.061)	115	Cross section
7		Latrine	0.219	0.155**	(0.040)	169	Cross section
8		Traditional midwife post	0.399	0.002	(0.106)	70	Cross section
9		Community center	0.251	0.244**	(0.053)	95	Cross section
	Panel C: Hypothesis 2 - Full Sample Outcomes						
10	Existence of functional local public good in the community:	Primary School	0.462	-0.007	(0.050)	464	Panel
11		Grain drying floor	0.237	0.104	(0.066)	459	Panel
12		Traditional midwife post	0.079	0.175**	(0.035)	235	Cross section
13		Latrine	0.462	0.210**	(0.059)	234	Cross section
14		Community center	0.212	0.241**	(0.063)	469	Panel
15	Community took a proposal to an NGO or donor for funding		0.292	-0.156+	(0.081)	460	Panel
	Panel D: Hypothesis 2 - Conditional Outcomes						
16	Supervisor's physical assessment of construction quality (index from 0 to 1):	Primary School	0.583	0.116*	(0.055)	123	Cross section
17		Grain drying floor	0.375	0.142+	(0.076)	101	Cross section
18		Latrine	0.270	0.177**	(0.055)	154	Cross section
22	Money and supplies were provided at least in part by the community:	Primary School	0.554	-0.007	(0.112)	242	Panel
23		Grain drying floor	0.105	0.086	(0.124)	184	Panel
24		Latrine	0.761	-0.197*	(0.093)	126	Cross section
	Panel E: Hypothesis 3 - Full Sample Outcomes						
25	Total petty traders in village		2.432	0.719*	(0.344)	225	Cross section
26	Total goods on sale of 10		4.449	0.560*	(0.240)	236	Cross section
27	Household asset score		-0.170	0.212*	(0.090)	471	Panel
28	Household asset quintile		2.835	0.158+	(0.094)	471	Panel
29	Attended skills training		0.061	0.119**	(0.018)	235	Cross section
30	Income from top 3 cash earning sources (in 1,000 Leones)		746.94	-21.773	(73.069)	236	Cross section

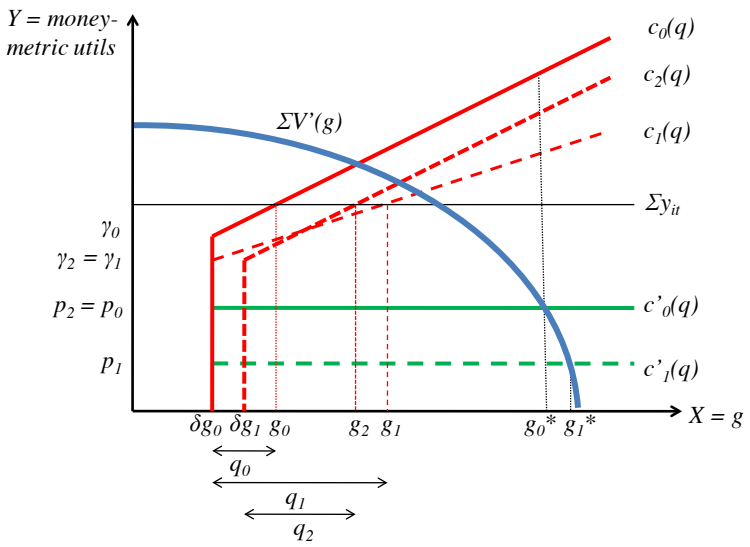
Notes on table: i) significance levels indicated by + p <0.10, * p <0.05, ** p <0.01; ii) the relevant treatment effect for the panel specifications is the coefficient on GoBifo*Post, while the relevant treatment effect for the cross section (follow-up data only) specifications is on GoBifo; iii) robust standard errors in parentheses, clustered by village for panel specifications; iv) includes fixed effects for the district council wards (the unit of stratification); v) 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; and vi) "conditional" outcomes are conditioned on the existence of a public good.

Table 4: Illustrative Treatment Effect Estimates for the Structured Community Activities (SCAs)

Structured Community Activity (SCA) Outcome:	Mean for Controls	Treatment Effect	Standard Error	N
	(1)	(2)	(3)	(5)
Panel A. Collective Action and the Building Materials Vouchers				
GoBifo Mean Effect for SCA #1 (13 outcomes in total)	**0.00**	**-0.057**	**(0.053)**	**236**
Proportion of communities that redeemed any vouchers at the building material supply store	0.54	-0.01	(0.06)	236
Average number of vouchers redeemed at the store (out of six)	2.95	0.11	(0.35)	236
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**
Duration of gift choice deliberation (in minutes)	9.36	1.60	(1.13)	225
Total adults in attendance at gift choice meeting	54.51	3.50	(3.20)	236
Total women in attendance at gift choice meeting	24.99	1.99	(1.68)	236
Total youths (approximately 18 to 35 years old) in attendance at gift choice meeting	23.57	2.10	(1.38)	236
Total number of public speakers during the deliberation	6.04	0.24	(0.40)	236
Total number of women who spoke publicly during the deliberation	1.88	-0.19	(0.22)	236
Total number of youths (approximately 18 to 35 years old) who spoke publicly	2.14	0.23	(0.24)	236
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**
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
Proportion of communities that stored the tarp in a public place	0.06	0.06	(0.04)	225
Proportion of communities that had used the tarp by the follow up visit (5 months after receipt)	0.90	-0.08+	(0.04)	233
Given use of the tarp, proportion of communities that used the tarp in a public way	0.86	0.02	(0.05)	161

Notes on table: i) significance levels denoted by + $p < 0.10$, * $p < 0.05$ and ** $p < 0.01$; ii) robust standard errors; iii) treatment effects estimated on follow-up data; (iv) includes fixed effects for the district council wards (the unit of stratification); and v) 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.

Figure 1: The Social Planner's Optimization Problem Facing a Constrained Budget



The International Growth Centre (IGC) aims to promote sustainable growth in developing countries by providing demand-led policy advice based on frontier research.

Find out more about
our work on our website
www.theigc.org

For media or communications
enquiries, please contact
mail@theigc.org

Subscribe to our newsletter
and topic updates
www.theigc.org/newsletter

Follow us on Twitter
@the_igc

Contact us
International Growth Centre,
London School of Economic
and Political Science,
Houghton Street,
London WC2A 2AE

IGC
International
Growth Centre

DIRECTED BY
LSE
UNIVERSITY OF OXFORD

FUNDED BY
UKaid
from the British people

Designed by soapbox.co.uk