A Critical Review of Community-Driven Development Programmes in Conflict-Affected Contexts

Elisabeth King, Ph.D.
Balsille School of International Affairs
eking@balsillieschool.ca

**A Critical Review of Community-Driven Development Programmes
in Conflict-Affected Contexts**

Contents

Executive Summary

- After participating in two rigorous impact evaluations of Community-Driven Development/Reconstruction (CDD/R) in Liberia and DRC, IRC and DFID embarked on this review as a next step in learning. They also wanted this review to inform design and evaluation strategies for new CDR programming in Somalia.

- CDD/R programmes – that empower local communities to directly participate in development activities and to control resources to do so – aim to improve socio-economic wellbeing, governance, and social cohesion at a local level. While CDD/R is context-driven, it is generally implemented as a standard model.

- According to rigorous impact evaluations from programmes in Afghanistan, Democratic Republic of Congo (DRC), Aceh (Indonesia), Liberia and Sierra Leone, and interviews with practitioners, policymakers and academics, the record of CDD/R in conflict-affected contexts is mixed and, overall, disappointing in terms of reaching the ambitious goals set out.

- As currently designed, implemented, and evaluated, CDD/R is better at generating the more tangible economic outcomes than it is at generating social changes related to governance and social cohesion, although even the economic effects are found in just a few studies. Moreover, CDD/R programming is better at producing outcomes directly associated with the project rather than broader changes in routine life.

- CDD/R has been plagued by a panacea-type approach to goals and a generalised theory of change that is, as interviewees characterised it, "lofty", "unrealistic", "inherently flawed" and even "ridiculous".

- A variety of issues related to programme design merit rethinking: the relatively short timeline of CDD/R projects, the small size of block grants, the limited reach of the projects, the menu restrictions on CDD/R programming, the limitations of social infrastructure, the quality and intensity of social facilitation, the manner in which communities are conceptualised and thus often not meaningful to participants, and how community institutions build on existing institutions and relate to the state.

- Although the evaluations reviewed here are of high quality, they raise a number of methodological questions about the best measures and instruments for evaluating CDD/R, the timing of measurement, and levels of analysis, as well as if and how evaluations impact projects and outcomes.

- Open and honest conversation about CDD/R – which has occurred too infrequently – must guide the way forward.

- Future CDD/R efforts also need to be guided by humility and more realistic goals.

- More questions can and should be asked in evaluations. Areas for future research on CDD/R consist of comparing CDD/R to other programming rather than a counterfactual of no programme, parsing the social and economic aspects of programme inputs and consequent outcomes, introducing variation within treatment communities to learn more about programme design and contextual features, and asking how and why questions

about the CDD/R process, and the outcomes it generates. Stronger monitoring is essential.

- The road ahead must build on the important work undertaken so far and the many questions raised here, not simply replicate what has been done in the past.

Acknowledgements

Introduction and Motivation

The International Rescue Committee (IRC) partnered with academics from Columbia and Stanford Universities to conduct rigorous impact evaluations of its Department for International Development (UK)-funded CDD/R projects in the conflict-affected contexts of Liberia and DRC.[1] Both evaluations were forerunners in terms of their identification and measurement strategies and ultimately, their ability to detect impacts. Neither these evaluations, nor the others studied here, were typical commissioned evaluations; each evaluation was designed and conducted by an academic team with much freedom. In Liberia, the evaluation team found "powerful evidence that the programme was successful in increasing social cohesion" – a finding which stands out from other evaluations of CDD programming – paired with weak evidence of changes to governance and economic well-being (Fearon, Humphreys, and Weinstein 2008 i). The subsequent evaluation of the Tuungane programme, implemented in eastern DRC, showed that despite the high quality of implementation, and "…little evidence of adverse effects, the evidence for positive effects [on economic well-being and socio-political attitudes and behaviors] is scattered and generally weak" (Humphreys, Sanchez de la Sierra, and van der Windt 2012, 8). The authors of the evaluation conclude that "coupled with negative evidence from related studies our findings present a challenge to the CDR model and its ability to produce the social and economic impacts that advocates attribute to it" (ibid). The IRC and DFID embarked on this review as a next step in learning. Meanwhile, DFID and IRC are working with partners to design a CDD-related programme in Somalia; they also wanted this review to inform programme design and evaluation strategies.

The timeliness of this review of CDD/R is also supported by the state of the broader literature. References abound to the amount of money spent on CDD/R[2] and the promise of CDD/R remains widely cited. A number of rigorous evaluations have emerged since Mansuri and Rao (2004) wrote that evidence on CDD "lags well behind the rate at which projects are being implemented and scaled up". But, results are mixed, making synthetic review timely. Two such reviews emerge from the World Bank: Mansuri and Rao (2012) conduct an extensive review of "induced" participation and Wong (2012) studies CDD specifically, examining World Bank-funded projects evaluated to rigorous impact evaluation standards. This review distinguishes itself from these other efforts in several respects: it focuses more narrowly on CDD/R in conflict and post-conflict contexts evaluated to rigorous impact evaluation standards, includes projects beyond the World Bank, and frames findings and analysis in terms of a theory of change. In addition, the questions for this review are specifically generated by IRC and DFID at this key juncture in decision-making.

---

[1] By impact evaluation, we mean "a 'with versus without' analysis: what happened with the programme (a factual record) compared to what would have happened in the absence of the programme (which requires a counterfactual, either implicit or explicit)… to attribute some part of observed changes to the policy, programme or project being evaluated" (White 2013).

[2] For example, Wong notes that the World Bank currently supports over 400 CDD projects in 94 countries, valued at over $30 billion (2012, iv).

Indeed, the list of questions guiding this review is long (see technical appendix for a full list) and the stakeholders are diverse. The list includes identifying and discussing a theory of change; reporting on main findings from rigorous impact evaluations to date; considering the factors that explain the mixed results; identifying practical lessons for evaluations and programming; suggesting alternate hypotheses emerging from existing studies; and identifying the next questions for researchers. Given that each of these questions and the many related sub-questions could form the basis of an entire report, the review endeavours to balance among these priorities. In the end, it reports findings while simultaneously raising questions and aims to open honest debate about CDD/R programming and evaluation.

This review proceeds in six parts. The first provides background including the definition of CDD/R, reflections on conflict-affected contexts in light of CDD/R, and the goals of CDD/R. The second part details the methods of this review. The third part presents the theory of change. The fourth part assesses study designs and synthesizes results from rigorous evaluations of CDD/R in Afghanistan, DRC, Aceh (Indonesia), Liberia and Sierra Leone. The fifth part explores multiple possible explanations for the results, ranging from programme goals and theory to programme inputs and outputs, and from programme design and implementation to evaluation strategy. The final part makes suggestions for the way forward for policy-makers, practitioners, and researchers working on CDD/R.

Background

What is CDD/R?

CDD/R is defined as an "approach that empowers local community groups, including local government, by giving direct *control* to the community over *planning decisions* and *investment resources* through a process that emphasizes participatory planning and accountability" (Social Development Department 2006, emphasis added). The approach is thought to be "a way to provide social and infrastructure services, organize economic activity and resource management, empower poor people, improve governance, and enhance security of the poorest" (Dongier et al. 2001).

Indeed, the very names of the projects speak to the definition and goals of CDD/R: *Tuungane* (DRC) means "let's unite" in Swahili, *GoBifo* (Sierra Leone) means "move forward" or "forward march" in Krio, and Afghanistan's CDD/R programme is called the National Solidarity Programme.

It is oft-repeated that CDD is "context driven" (i.e. World Bank 2006, 41). In interviews, this flexibility was often touted as one of the principal strengths of the CDD/R approach. Nonetheless, what some call a relatively "standard model" underlies the principal evaluations we

review here. These, and most other, CDD/R projects have two primary components: (1) a community- or institution-building and planning component, usually including the election of local community councils; and (2) a block grant for a project, or "asset investment component".

Conflict-affected contexts

We focus on CDD/R interventions implemented in conflict or post-conflict communities. There are a number of assumptions about conflict-affected communities that make CDD/R and its goals potentially particularly fitting (McBride and D'Onofrio 2008; McBride and Patel 2007; Barron 2010; Cliffe, Guggenheim, and Kostner 2003). First, there are presumptions about the consequences of conflict. In terms of socio-economic factors, physical infrastructure may be destroyed. Supply chains and service provisions are likely to be disrupted and, in cases where services do exist, people are likely to encounter obstacles to access. There may be a need for a tangible "peace dividend". In terms of governance, conflict-affected societies are characterized by weak or absent institutions, lack of experience with good governance, and weak state-society relations. In terms of social cohesion, assumptions include a breakdown in social cohesion, inability or unwillingness of communities to work together and lack of trust among community members and towards institutions.[3] Women and youth may be disproportionately disadvantaged, alongside other vulnerable groups such as ex-combatants and widows. These types of conceptual presumptions are repeated in much of the project-specific documentation (i.e. see Box 1). The World Bank CDD website lists "ten reasons why the CDD approach makes sense in a conflict or post-conflict setting." [4]

Box 1: Contextual Presumptions in Liberia

"Fourteen years of intermittent conflict and over **100 years of corruption and exclusion have devastated the political, economic, and social fabric of Liberia**….With its **roots in bad governance** as well as **marginalization …**conflict has broken community and familial relationships and laid waste to the trust in institutions deemed essential to the recovery process. It has further **destroyed physical infrastructure** and rendered the systems, structures, and mechanisms for delivery of social services almost non-existent."

(IRC 2006c, 1-2, emphasis added)

There are also presumptions about the roots of conflicts. For example, the CDD/R approach works on the premise that at least some of the roots of conflict are at the local community level (USAID 2007, 3) and relate to exclusion and inequity (World Bank 2006). In Sierra Leone, for instance, the governance goals of the CDD/R project aimed to address some of the professed root

---

[3] Mansuri and Rao call this situation a "civil society failure at the local level"; a "situation in which groups that live in geographic proximity are unable to act collectively to reach a feasible and preferable outcome" (2012, 4).
[4] See:
http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTSOCIALDEVELOPMENT/EXTCDD/0,,contentM
DK:23006650~menuPK:549351~pagePK:148956~piPK:216618~theSitePK:430161,00.html

causes of the war (Casey, Glennerster, and Miguel 2011a, 3). Similarly, the idea that poverty, poor local level governance, and marginalisation were root causes of conflict informed the CDD/R project in Liberia (Fearon, Humphreys, and Weinstein 2011 8). In the academic peace and conflict literature, as well as policy literature, there is an important trend towards recognizing the importance of local factors in conflict (i.e. Autesserre 2010; Kalyvas 2006; Straus 2006; Dennys and Zaman 2009) and in peacebuilding (Hussein 2004; Bouvy and Lange 2012; Autesserre 2010; Peace Direct and Quaker United Nations Office 2010; Leonard 2013).

Alongside these high needs, conflict-affected communities may also face a window of opportunity for change and progress. The World Bank refers to "new 'development spaces'" which arise as conflicts unsettle the status quo (2006, p.12). Project documentation also refers to special opportunities in conflict affected states. Eastern DRC is described in early Tuungane documents as a situation "of huge suffering but also huge potential" (IRC 2006a, 11) and project literature from Liberia notes the chance to "seize this window of opportunity to re-enforce the peace" (IRC 2006b, 1).

It is oft-repeated that the CDD/R approach can be used at different stages of conflict: for prevention, during conflict, in preparation for peace processes, as part of post-conflict recovery, and during post-conflict peacebuilding and development (Haider 2009 5; World Bank 2006, 9, 46). Indeed, the country cases we study here are implemented in some of the most difficult conflict-affected contexts in the world. Conflict is ongoing in DRC and Afghanistan; these two states are near the top of the failed state index and the bottom of the human development index (Foreign Policy and Fund for Peace 2012; UNDP 2011).[5] The CDD/R programmes in Liberia, Sierra Leone and Aceh also faced difficult challenges: each was implemented in a different post-conflict context. On the human development index, Liberia ranks 182nd and Sierra Leone ranks 180th (of 187), both in the low human development category, while Indonesia, not regionally disaggregated, ranks 124th (UNDP 2011).

In response to these contexts, CDD/R interventions may integrate a specific reconstruction or reconciliatory component which differs from CDD in non-conflict-affected communities, but they may not. The Sierra Leone study notes, for instance, that the design of GoBifo is similar to the design of CDD in non-conflict affected contexts (Casey, Glennerster, and Miguel 2011a, 3).

Goals of CDD/R

According to the literature on CDD/R, there are three principal goals of CDD/R interventions: improved socio-economic recovery; improved social cohesion; and improved governance (Social Development Department 2006; International Rescue Committee undated). Of the five key

---

[5] DRC and Afghanistan rank second and sixth respectively on the Failed States Index and rank 187th and 172nd respectively, out of 187, on the Human Development Index.

studies under review, all mention economic, governance, and social goals (see Table 1). In most conflict-affected cases, all of these goals of CDD/R relate in some way, moreover, to building peace and enhancing security (World Bank 2006).

Figure 1: Project Goals

Project	Goals as per project documentation and/or evaluations
Afghanistan	"The key objective of NSP is to build, strengthen and maintain Community Development Councils (CDCs) as effective institutions for local **governance** and **social-economic development**" (http://nspafghanistan.org/default.aspx?sel=26#Q6). Beath et al. add that the program "explicitly mentions promoting **gender equality** as one of the programme's main goals" (Beath, Christia, and Enikolopov 2012b 6).
Aceh (Indonesia)	"The programme had both **economic** and **social goals**. It aimed to deliver quick assistance to conflict-affected villagers in order to improve material wellbeing in the short term. In addition, it sought to promote social cohesion, strengthen village-level decision-making institutions, and cultivate **greater faith in governmental institutions** in the aftermath of the conflict." (Barron et al. 2009) "The primary focus of BRA-KDP is to assist conflict-affected communities in improving their living conditions through provision of small projects that accord with their needs. It also encourages people to **overcome mistrust of government that is a result of the conflict by delivering tangible outputs** that fit with communities' priorities" (Barron et al. 2009).
DRC	"To improve the **stability** and **quality of life** for communities in eastern DRC through structured, participatory, and inclusive collective action. By establishing and strengthening participatory local governance committees [the programme aims…] to improve the **understanding and practice of democratic governance**, improve citizens' relationships with local government, and **improve social cohesion** and thereby communities' ability to resolve conflict peacefully. The conduit to achieve these purposes will be village- and community-level projects that themselves will contribute to socio-economic rehabilitation as DRC moves into a post-conflict and development period" (Humphreys, Sanchez de la Sierra, and van der Windt 2012, 11). "While people's vision of democracy's dividends [in anticipation of Congolese elections] is in all probability unrealistic in time and scope, it is nevertheless vital that they receive **tangible returns** for their enduring tolerance. It is thus crucial that the post-election period **deliver peace dividends**…" (IRC 2006a, 11)
Liberia	"…the project aims to **improve material welfare**, **build institutions and promote community cohesion** by bringing together all actors within the community, including local government, civil society and private sector to identify priority problems/needs and to develop community action plans for implementation." (IRC 2006b, 1) "This model was adopted, in part, as a strategy for using local leadership to quickly generate material improvements in people's lives. Given the state of the government after fifteen years of civil war communities could plausibly also move more quickly than government to deliver a **tangible peace dividend**" (IRC).
Sierra Leone	"Through intensive, long term facilitation, CDD aims to **strengthen local institutions**, make them more democratic and inclusive of marginalized groups, and enhance the capacity of communities to engage in **collective action**" (Casey, Glennerster, and Miguel 2011c, 1).

Study hypotheses speak to these goals (see Box 2). In project documents, broader literature, as well as interviews, the idea that these goals are mutually reinforcing or interdependent emerges

frequently. In other words, as elaborated in the Theory of Change (ToC) below, people believe that attainment of one goal indirectly helps foster the attainment of the others.

Box 2: Pre-analysis Hypotheses from Tuungane, DRC

W1 Greater household income and asset holdings
W2 Greater share of time allocated to productive activities
W3 Improved performance on health and education indicators
G1 Greater willingness to engage in local decision-making processes
G2 Increased willingness to hold traditional and political leaders accountable
G3 Greater ability to implement development initiatives in a transparent and equitable way
S1 Improved social cohesion
E1 Better project implementation in parity areas
E2 Better outcomes for women in governance and welfare indicators in parity areas
E3 Greater propensity for women to take on leadership roles in parity areas

(Humphreys, Sanchez de la Sierra and van der Windt 2011, 6)

In conflict-affected contexts, the idea of "peace dividends", a shorter term goal parallel to infrastructure and service delivery goals in non-conflict affected contexts, emerges equally frequently (Cliffe, Guggenheim, and Kostner 2003; World Bank 2006, see again Figure 1). A peace dividend is a tangible benefit meant to illustrate the value of peace. CDD/R in conflict-affected contexts allows for the merging of crisis response and development (Cliffe, Guggenheim, and Kostner 2003, 2).

Alongside these commonalities and principally recognized goals, there are also some divergences in goals. Specific projects may make one or more goals paramount to others. For instance, an interview about Tuungane suggested that at least early on in the programme, the governance goal was paramount to others. Some projects emphasize additional goals. The NSP in Afghanistan promotes gender equality as one of the programme's principal goals. Furthermore, there are often non-official goals of CDD/R programmes in post-conflict contexts. For example, BRA-KDP in Aceh came in response to major problems with a prior method of assistance delivery to ex-combatants and victims of conflict. As a result, one of the aims of the CDD/R programme in Aceh was to distribute money in a way that minimized harm and that avoided the mistakes of the previous programme. Goals related to terrorism and counter-insurgency emerged in some interviews about the NSP in Afghanistan. Finally, it was common in interviews to hear different goals or goal priorities from different people on the same projects; there are disconnects between and among field staff, operational manuals, main office staff, and evaluators. Goals may also change over time as the project evolves and primary personnel turn over. Indeed, it appears that key individuals are influential in programme choices and goals.

Review Methods

This review relied on three principal methods. First, it is based on existing CDD and CDR literature, focusing especially on previous reviews and papers related to conflict-affected environments. References to these works appear throughout the report.

Second, in order to uncover the main findings of rigorous CDD/R evaluations to date, we searched the internet, and through databases and contacts for studies based on two inclusion criteria: (1) studies must focus on CDD/R as per the definition above and take place in a conflict-affected context and (2) studies must have clear research questions and sound evaluation designs that include estimation of a counterfactual.[6] These criteria resulted in a number of exclusions. We first excluded studies of CDD that did not focus on cases in conflict-affected contexts. The most difficult decisions were to exclude studies of the KDP/PNPM in Indonesia and the KALAHI programme in Philippines because, while each country has conflict-affected regions, and the studies meet evaluation standards, the results are not disaggregated by region, making it difficult to see if and how the findings hold in areas directly affected by conflict – those in which we are most interested. Next, we excluded studies of CDD that did not meet the research design standard we set. Since lack of impact evaluation of CDD/R has been a common critique in the literature (see i.e. Barron 2010, 28; Mansuri and Rao 2012, 13), this narrowed the pool quite substantially. In the end, this review includes five evaluations – stemming from CDD/R programmes in Afghanistan, DRC, Indonesia (Aceh), Liberia, and Sierra Leone.[7] It is quite remarkable that these studies are all large in scale, have relatively strong identification and measurement strategies, and study programmes with a number of common core elements. As such, we took studies with high internal validity and wanted to learn more about external validity. Based on the theory of change and overview documents, we created a reading template which helped us to systematically extract programme and evaluation features of each study and, to the extent possible, collect information in accordance with the theory of change (see technical appendix). Much of the analysis that follows is derived from these five studies.

Third, we conducted fifteen interviews with individuals with expertise on CDD/R in conflict-affected states. This included at least one author of each of the studies included here, as well as authors of previous reviews and practitioners from implementing agencies. Interviewees spoke in their personal capacities, rather than as representatives of their present or past organisations and all of their comments remain unattributed in the text below. While every interview continued to add rich nuance and analysis, after fifteen interviews, there was a general consensus on main issues or at least main lines of disagreement were clear. The discussion guide for the interviews

---

[6] The commissioning agencies themselves chose these criteria, acknowledging the pros and cons of restricting the body of evidence in this way. By counterfactual we mean that "for large-n impact evaluation designs, the counterfactual is constructed by identifying a comparison group, which is similar in all respects to those receiving the intervention, except that it does not receive the intervention." For small-n studies, estimating the counterfactual may employ "deductive approaches based around causal chain analysis, such as process tracing" (White 2013).

[7] This review went to press prior to the publication of the final NSP report on Afghanistan. We also expect that an evaluation from Sudan should follow relatively shortly.

11

is available in the technical appendix. Notes and transcripts from the interviews were analysed with Atlas-ti, a qualitative data analysis programme. Valuable insights from the interviews are integrated throughout the analysis.

How CDD/R is supposed to work: Theories of Change (ToC)

What is the underlying theory of change of the CDD/R model?

According to the International Initiative for Impact Evaluation (3ie) guide on impact evaluation process, a theory of change "clearly lay[s] out how it is that the intervention (inputs) is expected to affect final outcomes" (3ie n.d., 2). In this way, a CDD/R ToC should take the goals explained above, and try to map out the preconditions and pathways needed to achieve them. It should lay out principal assumptions, identify key indicators of progress and suggest a timeline over which effects are likely to transpire (Mansuri and Rao 2012, 288). Asking the question in this way presupposes the existence of one dominant ToC and one dominant model of CDD/R.

One of the clearest findings of this review, from both literature and interviews, is that there is no clearly articulated, theoretically well-founded, consistent, theory of change guiding CDD/R programming. There is, nonetheless, a generalised theory that several documents (International Development Association 2009; McBride and D'Onofrio 2008; Cliffe, Guggenheim, and Kostner 2003) and many interviewees repeat. It is summarized in Figure 2. The general idea is that through well-implemented CDD/R intervention inputs – that is mobilization, training, and bolstering or creating inclusive community-level institutions, followed by a block grant to be used for one or more projects chosen by the community – the following *outputs* are generated:

- People participate in decision-making. In so doing, they shift existing power arrangements.
- Through inclusive institutions and processes, divided people, or at least people with collective action problems, work together.
- The community determines investment priorities. This induces accountability, incentives to economize, and collective ownership. Projects reflect people's priorities.

The theory continues that, as a consequence of programme inputs and these outputs, CDD/R results in intermediate, proximate outcomes, or "quick wins": generally social infrastructure built with the block grants. This social infrastructure may also be considered a peace dividend, tangibly showing citizens the benefits of peace, and therefore an endpoint goal (IRC 2006a; Cliffe, Guggenheim, and Kostner 2003). Haider notes that local levels of bureaucracy may be able to deliver outcomes more quickly than centralized governments (2009, 10), and in some cases, such as Afghanistan and DRC, the central state's capacity to reach most rural communities is simply not there.

A second intermediate outcome for CDD/R programmes is empowerment, or "the exercise of voice and choice" (Mansuri and Rao 2012, 15), characterized by a situation in which "community members…belie[ve] that they can affect change and can improve their own lives" (Haider 2009, 9). Like social infrastructure, empowerment is both instrumental and of intrinsic value. That empowerment/voice has intrinsic value was a major finding of the large participatory research project, *Voices of the Poor*, which collected the views of 60,000 poor people in 60 countries. The analysis of these data found that poverty manifests itself in many non-material outcomes, such as feelings of powerlessness, lack of voice, exclusion, breakdown of the social fabric, dependency, and shame (Narayan et al. 2000). Nobel-prize winner Amartya Sen (1999) characterized "Development as Freedom" and the World Bank 2000/1 World Development Report picked up on this theme and included empowerment as one of the three parts of poverty reduction.

There are two main pathways through which these inputs, outputs, and intermediate outcomes may lead to final outcomes. First, they may lead *directly* to the final outcomes. For instance, new or repaired infrastructure such as a road may improve access to markets, health care and education, directly improving community members' socio-economic welfare.

Figure 2: Generalised Theory of Change
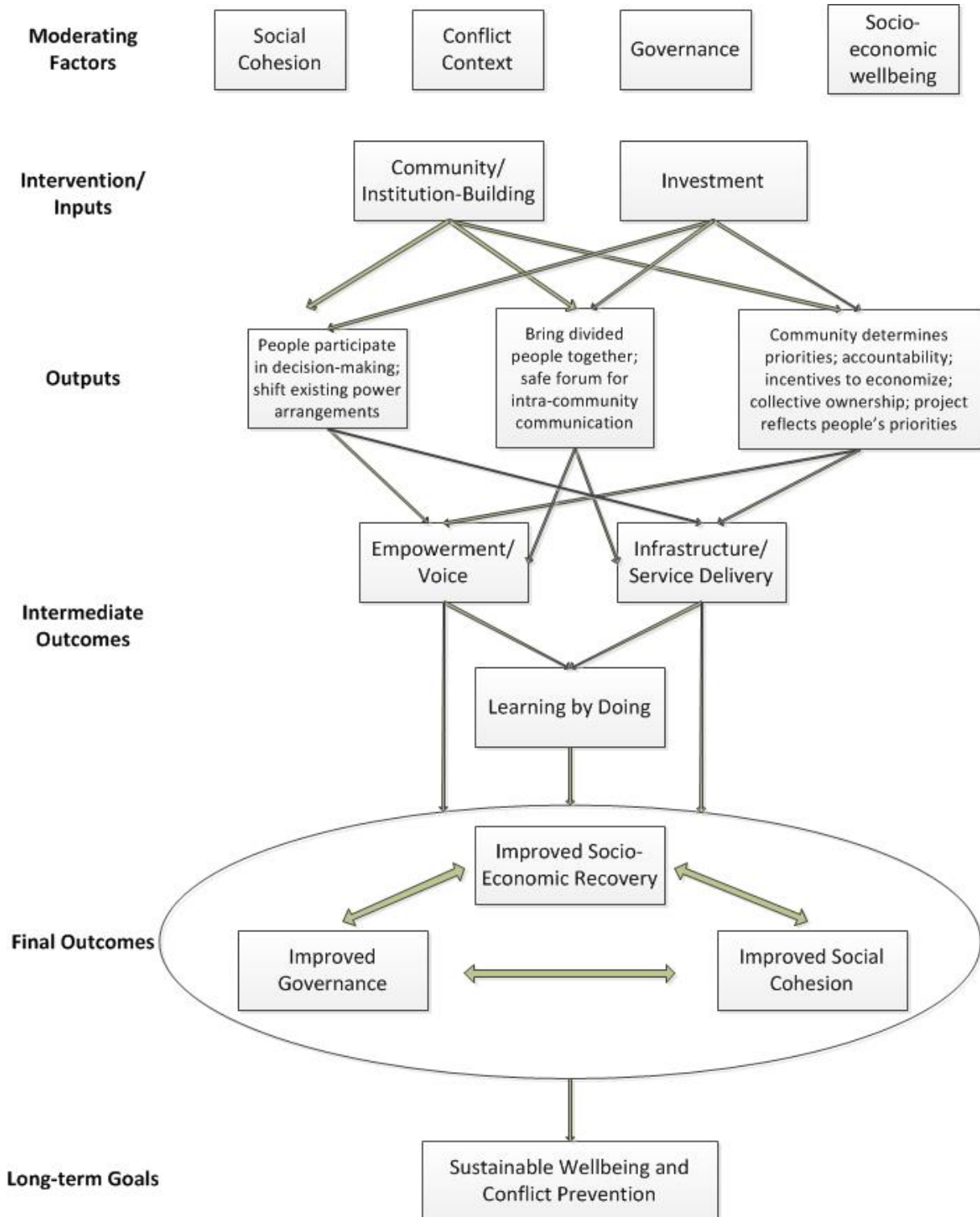
Second, alternatively or simultaneously, the proximate outcomes may lead *indirectly* to the final outcomes through "learning by doing", or what others call "demonstration effects". The idea is that community members see participatory programmes in action, see that they can affect processes in this way, see that a good product emerges – in other words they see utility in the process – and decide to try it out on further issues. For example, if, through the CDD/R experience, community members gain practice with cooperation and conflict management, they may effectively manage future conflicts through similar channels. Or, by engaging community members in decision-making, a precedent may be set for participatory decision-making, thereby increasing participation in governance. Or, as a final example, by seeing an example of successful collective action, community-members may engage in future collective action. Several of the included studies make mention of changes invoked through CDD/R programming carrying over into routine life as an ultimate goal.

Along either of these pathways, change may happen in regards to one or more final outcomes. Then, the logic is that the final outcomes are mutually reinforcing. For example, social cohesion may be strengthened, with the idea that through a heightened ability for collective action, community members will hold government leaders accountable, improving governance. They will also work together more, thereby improving welfare outcomes. While there are usually two-headed arrows envisioned between each of the final outcomes, the order – what is supposed to move first, or which factor is supposed to serve as the lever – is not specified.

In many of the post-conflict contexts, as noted above, it is posited that these ultimate goals could help build peace. For example, the Liberia evaluation notes, citing an earlier World Bank publication "by improving public goods provision or enhancing cohesion, CDR may reduce the risk of renewed conflict by lessening local grievances or facilitating economic development, which may in turn reduce the incentives to participate in violence" (Fearon, Humphreys, and Weinstein 2008, 4).

There are a number of moderating factors – meaning factors upon which the unfolding of this theory depends. As Barron writes, "where an environment is *conducive* to change, projects are more likely to act as a catalyst" (2010, 24). *Conflict context*, refers to the stage on a conflict-post-conflict-recovery continuum, degree of stability, and progress towards peacebuilding goals. It also refers to the numerous social, political, and economic assumptions, detailed above, that relate to conflict-affected contexts. As such, other important moderating factors include *pre-existing social cohesion,* referring to both vertical relationships between citizens and government and horizontal relationships between citizens; g*overnance* characterized by such factors as legitimacy, capabilities like spending capacity, support for decentralization (enhancing the likelihood that CDD structures and opportunities would be repeated), institutional history, and responsiveness; and *socio-economic wellbeing* including GDP/capita and access to resources. These factors support Mansuri and Rao's argument that challenges of local development arise at

the "intersection of market, government, and civil society failures" (2012, 4). While a strength of CDD/R is the ability to implement in difficult contexts, minimal security conditions is also a pre-condition for CDD/R (Cliffe, Guggenheim, and Kostner 2003, 5). All of these contextual factors matter at both local and higher levels (Mansuri and Rao 2012). We note the challenge that contextual factors and outcomes overlap.

Something along the lines of the generic theory of change that we present here is noted in most studies. See Figure 3.

Figure 3: Theory of Change in Project Documentation

Project	Theory of change as per project or evaluation
Afghanistan	"Mandated female participation can have powerful **demonstration effects**, signalling that women can be effective when taking part in village governance, and in turn reducing long held biases against them" (Beath, Christia, and Enikolopov 2012b, 11).
DRC	"The theory behind the Tuungane intervention is that training, coupled with exposure to and practice in accountable governance in the context of these projects, can produce **learning-by-doing** and bring about change in local accountability and social cohesion as well as improve the welfare of communities" (Humphreys, Sanchez de la Sierra, and van der Windt 2012). "Through various forms of capacity building, including project-based and intensive conflict-oriented programmatic work, communities, local authorities, and other stakeholders will increase their knowledge and their ability to work collaboratively. The resulting increase in both horizontal and vertical social capital will help to overcome mistrust and animosity, in addition to addressing the conflict-induced erosion of skills and resources. The gradual growth of the community's sense of self and ability to stand up for and realize its requirements will further its ability to pursue its needs and build a more equal relationship with local authorities"(IRC 2006a, 12).
Aceh (Indonesia)	"The primary focus of BRA-KDP is to assist conflict-affected communities in improving their living conditions through provision of small projects that accord with their needs. It also encourages people to overcome mistrust of government that is a result of the conflict **by delivering tangible outputs** that fit with **communities' priorities**. Equally important is the process by which **villagers identify, prioritize and implement** their projects." (Barron et al. 2009, 2)
Liberia	In the objective of "local experience in inclusive, transparent, and accountable decision-making (to achieve development goals)": "Communities participate in democratic election process… Local actors establish and adhere to internal operating procedures, rules, and norms… CDCs interrelate with government actors on development goals and activities… CDC decision making incorporates the concerns and conditions of vulnerable/marginalized/marginalized populations." In the objective of "Improved aspects of socio-economic conditions": "Communities determine priority needs that will most improve their local socio-economic conditions… CDCs select, design, and implement prioritized projects based on the CDP [community development plan]. In the objective of "increased trust, cooperation, and cohesion among community members": "Communities engage wide spectrum of stakeholders including socially-excluded and marginalized groups in project process… CDCs consider conditions, abilities, needs of vulnerable and marginalized members to increase their benefit from the projects… Community members accept diversity as a positive aspect of community" (IRC 2006b). "A "**learning-by-doing**" tactic tends to be the means by which CDR builds the capacity of CDCs, contractors and advisory groups." (IRC 2008, 11)

Sierra Leone	"CDD's more far reaching objectives involves **teaching communities how to help themselves** in the long run, by giving them the **experiences and skills** to independently undertake initiatives after the programme has ended." "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 that shifts social norms, this experience could trigger a persistent increase in their benefits of participation, sustainably raising participation levels into the post-programme period" (Casey, Glennerster, and Miguel 2011c, 12-3).

Does CDD/R work? Findings from rigorous impact evaluations

Intervention characteristics

We include five key cases in our review and synthesis of rigorous impact evaluations of CDD/R in conflict-affected contexts. See Figure 4 for a summary of some of the intervention characteristics.

Figure 4: Selected Programme Characteristics

Case	Implementer	Size	Principal Programme Design
Afghanistan, National Solidarity Programme (NSP)	Ministry of Rural Rehabilitation and Development via 28 NGOs (including IRC)	29,947 villages across 361 of 398 district in Afghanistan	Electing gender-balanced community development councils; block grants
DRC, Tuungane	IRC and CARE	1250 villages grouped into 560 communities in eastern DRC	Electing village development committees; block grants
Aceh, BRA-KDP	BRA and KDP (Indonesian government)	1,724 villages in 17 rural districts	Village meetings; block grants
Liberia	IRC	82 communities	Electing community development councils; block grant
Sierra Leone, GoBifo	Ministry of Internal Affairs, Local Government and Rural Development	236 villages	Electing community development councils; block grant

In Afghanistan, the National Solidarity Programme (NSP) is executed by the Ministry of Rural Rehabilitation and Development (MRRD) and designed and funded by the World Bank, Afghanistan Trust Fund and bilateral donors. Twenty eight NGOs, including the IRC, implemented this largest development project in Afghanistan in 29,947 villages across 361 of Afghanistan's 398 districts. The project entailed the election of gender-balanced community development councils (CDC) by secret ballot and universal suffrage, followed by block grants for village-level projects chosen by CDCs in consultation with communities. The randomized evaluation focused on the second phase of NSP to study 250 treatment and 250 control villages.

In DRC, *Tuungane* was funded by DFID and implemented by CARE and the IRC.[8] Tuungane worked in 1250 war-affected villages, targeting 1,780,000 beneficiaries in eastern DRC. The intervention involved organizing the election of village development committees (VDC) followed by training for VDC members and VDC sensitization of community members on issues related to "leadership, good governance, and social inclusion". VDCs subsequently worked with community members to choose and implement development projects. The randomized evaluation focused on the first phase of Tuungane, following 1250 villages grouped into 280 treatment and 280 control communities.

In Aceh, Indonesia, the Community-Based Reintegration Assistance for Conflict Victims Programme (BRA-KDP) was managed by the Aceh Peace-Reintegration Agency (BRA) and the Kecamatan Development Programme (KDP), with technical assistance and support from the World Bank and funding from the Government of Indonesia. BRA's Project Implementation Unit and the KDP Regional Management Unit implemented the programme. BRA-KDP provided block grants to communities with the intention that communities distribute the grants to projects presented by conflict victims. Villagers held at least four community meetings facilitated by KDP to decide how to spend the money. The communities themselves identified conflict victims, voted for proposal approval, and disbursed the money to projects with a very open menu that could be individual or public goods. Implementation teams elected by the communities oversaw project activities. This evaluation, the only evaluation of a non-randomized intervention included here, was based on post facto matching of 67 treatment communities to 67 comparable control communities, paired with weighting and correcting for imbalances.

In Liberia, DFID funded the IRC to implement a Community Driven Reconstruction project. IRC conducted sensitization to the new development project including meetings with chiefs and elders and oversaw the election of community development councils (CDCs). Then, CDCs were empowered to oversee the selection and implementation of community-chosen development projects using a smaller, then larger block grant. The randomized evaluation studied 42 treatment and 41 comparison communities.

In Sierra Leone, the GoBifo CDD/R project was implemented by the Ministry of Internal Affairs, Local Government and Rural Development of Sierra Leone, with grants from the Japan Social Development Fund and Italian Trust Fund administered by the World Bank. International consultants from the Royal Tropical Institute (KIT) helped with design. GoBifo was intended as a pilot for Sierra Leone's broader decentralization and participatory democracy agenda, and thus implemented in one area in the north, and a second in the south, to vary politics and ethnicity. GoBifo provided "technical assistance" that promoted such issues as democracy, the inclusion of

---

[8] IFESH was also involved in some implementation.

marginalised women and youth, and transparent budgeting. Communities were given the opportunity to manage block grants. The randomized evaluation included 118 treatment and 118 control communities.

While there are broad similarities among programmes, there are also important differences. These include, but are not limited to, the planned timeline (for example 18 months in Liberia vs. approximately 3 years in Afghanistan), who implements (government vs. NGOs, for example), the intensity of intervention (at least 4 community meetings in Aceh vs. roughly 6 months of direct facilitation over 3 years in each community in Sierra Leone) as well as its content, the number of staff working directly on CDR (18 in Liberia vs. upwards of 13,000 in Aceh), the number of intended direct beneficiaries, and the budget. Other sources of difference include the size of the block grants, the specifics of the negative list of projects (i.e. projects the programme will not fund), who exactly controls the money, whether and how much the community itself contributes to financing, the way that projects are selected, who implements the programme, whether the programme is implemented with new or existing community institutions, etc. The specific design features sometimes changed over the course of the programme. Many of these features are further discussed in the "is it programme design?" section below. We tried to compile a checklist of different components of programme design from evaluation and programme documents (see technical appendix reading instrument), but individual interviewees consistently corrected our classifications, often in contradictory ways.

Additionally, these descriptions suggest that there was a consistent programme across target areas. Indeed, each programme was designed to be implemented in roughly the same way in each of the included communities. But, there can also be tremendous variation in implementation at the local level. In Afghanistan, for instance, the 28 implementing NGOs had much discretion on programming in their respective areas (Maynard 2007). There was also variation over time in programming; the NSP, for example, has already had six different operational manuals with often quite significant differences from one to the next. In many contexts, there are also differences within countries as to when subprojects were completed (IRC 2008).

Study designs
Impact evaluations seek to answer the question "how much better or worse are programme areas than they would have been if the programme had not taken place?" In order to do so, they construct a counterfactual to ask "what would have happened if the programme had not taken place?" While they are not appropriate for every evaluation – the programme, the question one wants to answer, as well as practical and ethical considerations, must guide research methodology – it is widely recognized that the strongest strategy to estimate this counterfactual is an evaluation based on a randomised intervention (Karlan and Appel 2011; Banerjee and Duflo 2011; White 2013). In CDD/R, this works by indicating a pool of communities that are eligible to receive the project that is twice as large as the number that will ultimately receive the

project, then, randomly choosing, often through a public lottery, those that will receive the project.[9] In theory, unchosen communities (the controls), will be, on average, in all ways similar to those that are chosen (the treatment communities), except for the project. It is further presumed that the treatment communities would have followed the same trajectory as the controls, if not for the project, allowing researchers to draw conclusions about the impact of the project. As described above, four of the five studies included here, all except the Aceh study, are randomised interventions. Figure 5 summarizes some of the key features of these evaluations including assignment to treatment, principal instruments, approximate programme dates as compared to evaluation timing, and sample size.

Figure 5: Select Evaluation Design Features

Case	Assignment to Treatment	Principal Instruments	Approximate Programme Dates[10]	Timing of Main Data Collection	Sample Size[11]
Afghanistan, National Solidarity Programme (NSP)	Randomized	Baseline, midline, and endline surveys including questionnaires for male and female focus groups and male and female household/individuals; behavioural community activity.	Sept-Oct 2007 – Sept/Oct 2010	Baseline: Aug-Sept. 2007; Midline: May to Oct. 2009; Endline: May to Oct. 2011	250 treatment and 250 comparison villages
DRC, Tuungane	Randomized	Household Survey; Chief Survey; various key informant surveys (i.e. chief, enumerator, etc.); Behavioural data measures from unconditional cash transfer intervention	July 2007-June 2010	Baseline: July-August 2007; Endline and cash transfer in select communities: August 2010	280 treatment and 280 comparison communities
Aceh, BRA-KDP	Assignment via a scoring method based on conflict history and local institutional capacity	Household and village head surveys; individual behavioural measure.	August 2006-August 2007	Endline: July-September 2008	67 treatment, 67 control sub-districts
Liberia	Randomized	Household and chief surveys; behavioural games.	Nov 2006-March 2008	Baseline: March-April 2006; Endline: March-April 2008; Post-endline: July-September 2008	41 treatment and 41 control communities

[9] Sometimes, the pool of communities for the evaluation is narrower than the pool of communities that receive the project. By comparing Figures 4 and 5, for instance, the NSP was implemented in 29,947 villages across 361 of 398 districts in Afghanistan and the evaluation covered 250 treatment and 250 comparison communities.
[10] Note that there is some variation from community to community within a programme and sometimes slightly different reporting from different documents and individuals for each programme.
[11] See technical appendix for sample size for different measures.

Sierra Leone	Randomized	Household surveys; village-level focus groups; structured community activities (SCAs).	Jan 2006-July 2009	Baseline: Oct 2005; Endline: May-June 2009; post-endline: Oct. 2009.	118 treatment and 118 control communities

In terms of measurement, there is no standard, or agreed upon, set of indicators for concepts like social cohesion and good governance. This presents challenges for synthesizing results across studies. Key concepts, such as participation and community, may be differently defined in diverse context (Mansuri and Rao 2004, 31). When concepts are relatively agreed upon, such as "trust", different studies still measure it quite differently, referring for instance to generalised trust or trust in leaders (King, Samii, and Snilstveit 2010). They also use different questions. As Casey et al. write, "asking someone how much they 'trust' other members of their community is less likely to elicit meaningful responses than asking whether they have ever given money to a neighbor to buy something on their behalf at the local market" (Casey, Glennerster, and Miguel 2011b, 20). Even for economic wellbeing, where one might expect measures like income and consumption to be standard, some studies do not include these and different studies think about them differently.

Among these evaluations, we see a shift away from exclusively survey-based measures which characterised other evaluations of CDD/R, to the inclusion of behavioural measures via games and real-life activities that measure behaviour outside of the intervention. The inclusion of behavioural measures alongside attitudinal measures is considered a best practice since participants may tell researchers what they think they want to hear and, even when accurately reported, attitudes do always translate into behaviour (King, Samii, and Snilstveit 2010). Figure 6 shows how the studies included here made use of behavioural measures. These included significant innovations.

Figure 6: Behavioural Measures

Case	Targeted behaviour	Type and description of behavioural measure
Afghanistan	1) the quality of the targeting—whether aid recipients were among the vulnerable members of the community; 2) the extent of corruption—whether village leaders retained aid for themselves or for their relatives.	A distribution of wheat was organized to assess how NSP affects the quality of local governance. Village leaders in 491 communities in the sample received wheat from the World Food Programme to distribute it to the needy. To get at the mechanisms behind the effect of local institutions, there was a randomized variation in whether the wheat was handed out for distribution to people who are considered village leaders by the villagers themselves or whether elected officials or women are explicitly required to participate in the process (personal communication).
DRC	"Did areas that took part in *Tuungane* engage differently with RAPID relative to those that did not?" (Humphreys, Sanchez de la Sierra, and van der	A new intervention called RAPID was designed to assess the impact of the CDD/R programme. 560 communities (half of which participated in Tuungane and the other half did not) were selected to participate in an unconditional cash transfer programme in which they would receive grants of $1,000 to be used on village projects. Communities were actually told that they would be receiving $900, but $1,000 was in fact

	Windt 2012, 7)	given in order to provide a measurement of whether leaders report unanticipated gains to populations. There were no guidance as to who should manage the funds and how decisions should be made (Humphreys, Sanchez de la Sierra, and van der Windt 29-31)
Aceh (Indonesia)	"Trust in the ability or willingness of local gov't to use funds well…" (Barron et al. 2009, 58)	During the survey, in a behavioural public goods game, enumerators gave respondents about 1USD and asked participants to decide in private how much they wanted to contribute to development activities by the district and how much they wanted to keep for themselves (Barron et al. 2009 58).
Liberia	"Observe how communities conduct themselves when they confront major decisions and to what extent they are able to mobilize participation from community members" (Fearon, Humphreys, and Weinstein 2008, 25)	A community wide-public goods game, entirely separate from project implementers and evaluators, was announced one week in advance to communities. On game day, individuals were then selected at random from 24 randomly selected households, given approximately 5USD, and asked to privately decide how much to contribute to the community and how much to keep for themselves. There were also variants in this public goods game in which women only groups played the game, and in which there were variations in the amount of interest on community contributions (Fearon, Humphreys, and Weinstein 2008, 25; Fearon, Humphreys, and Weinstein 2011).
Sierra Leone	Respectively, (1) Ability for collective action; (2) participation, quality of participation; (3) elite capture and collective action.	Three "structured community activities" provided communities an asset and an opportunity, allowing the study team to observe how communities responded. These entailed (1) vouchers to each community that could only be redeemed if matching funds were raised; (2) community choice of one gift among two options; (3) gift of one tarp to each community (Casey, Glennerster, and Miguel 2011, 21-2)

When many outcomes are reported in testing a single hypothesis, there is the potential for analysts to incorrectly report "significant" results. That is, as Casey et al. write in the case of Sierra Leone, "given the large number of outcomes under consideration: [in their case] treatment effects for approximately 16 of the 318 unique outcomes would be statistically significant due purely to random choice" (Casey, Glennerster, and Miguel 2011b, 23). As Fearon et al. discuss, while a standard statistical approach may be to look at each outcome separately, pure chance would result in approximately 5% of estimates arising as statistically significant (Fearon, Humphreys, and Weinstein 2008, 36). Methods for avoiding such inferential biases include constructing "family indices" or using *p*-value adjustment methods, as was done in the Afghanistan, DRC, and Sierra Leone studies. Latter write-ups by the authors on the Liberia study (i.e. Fearon, Humphreys, and Weinstein 2011) also index results, which they did not do in the original impact evaluation report.

Some of these evaluations also introduce a new best practice in CDD/R literature: pre-analysis plans, in which hypotheses and specific indicators are specified prior to data collection and

analysis.[12] The researchers studying Afghanistan, DRC and Sierra Leone "bound their own hands" against cherry picking (selective presentation of empirical results) or data mining (in the sense of sifting through data to find significant results), which would have weakened their analysis. Moreover, without a pre-analysis plan, the authors could generate quite different interpretations of impacts, as the Sierra Leone illustrates through a thinking and data mining exercise (Casey, Glennerster, and Miguel 2011a). Crafting a pre-analysis plan is a useful thinking exercise for all involved in CDD/R. While there is some disagreement as to how much one can or should deviate from a pre-analysis plan in order to generate new hypotheses and uncover unexpected findings, pre-analysis plans generate systematic thinking and transparency. Overall, these studies demonstrate that it is possible to carry out innovative, rigorous impact evaluations in very difficult contexts.

What are the main findings of rigorous CDR/CDD evaluations to date?

Synthetic reviews aim to synthesize results across several studies and to thereby teach us more by pooling studies together than by regarding them individually. It is based on the idea that the whole is more than the sum of their parts. Statistical meta-analysis, where results are combined from different studies, would have been useful in theory, but in practice highly limited because of too many different measures for each outcome of interest across this quite small set of cases. We were unsatisfied with the idea of selecting outcome measures based only on repetition across studies rather than a priori relevance and, in practice, even selecting on repetition would have led to spotty results since so few measures are repeated (King, Samii, and Snilstveit 2010). There were also too few cases for a correlation-based analysis of outcomes with programme features or context. In other words, it would be useful to know if, for instance, larger block grants correlated with stronger outcomes, or if CDD/R implemented during conflict or after conflict worked better, but we did not have enough studies to do more than draw out hypotheses. Meta-analysis should be on the to-do list for the future, once the studies are published and if the study authors were willing to share either the raw data, or summary statistics plus covariance details, for key variables. More correlation-based analysis will be possible as more studies emerge, or possibly, by expanding the pool of included studies, an issue we discuss below. At this time, we present check-box type synthesis of findings across studies as well as detailed statistical tables from each study. We recognize that check-box type of synthetic review has its limits, but that it is also useful as a way to think across studies and to see patterns (3IE 2012).

Indeed, there are a number of different ways to think about the principal findings and – importantly – different modes of thinking about them lead to different conclusions. First, we can report the main findings as they relate to the final outcomes of interest taking the authors' own classifications of their measures and their descriptions of their results. Figure 7 borrows the colours of a stop light to illustrate statistically significant positive findings (green), indeterminate findings (amber), and lack of positive effects (red). We include detailed findings from each study

---

[12] There is also a move towards registering pre-analysis plans (White 2013).

in the technical appendix. In this appendix, we report the indicator, control mean, treatment effect, standard error, t-statistic, sample and source for each of the five principal studies. Since the treatment effect focuses on differences between treatment and control communities, we also include the control mean at endline to provide contextual information and interpretation of substantive effects, although this is not possible for standardized index measures. The smaller the standard errors in relation to treatment effect, the more confident we can be in the result. The control mean shows the counterfactual end state against which we can compare outcomes in programme locations. We have included t-statistics as an indication of significance.

Figure 7: Findings on Key Outcomes from Five Impact Evaluations

Outcomes		Afghanistan	DRC	Aceh (Indonesia)	Liberia	Sierra Leone
Proximate	Peace dividend (social infrastructure, service delivery, short-run economic)		yellow	red	red	green
	Empowerment/voice	green	red	red	yellow	yellow
Final outcomes (indirect)	Improved economic wellbeing	yellow	red	yellow	red	green
	Improved governance	yellow	red	red	yellow	yellow
	Improved social cohesion		red	red	green	red

In synthesizing these results, a mixed picture of results emerges. In terms of economic welfare, the study in Sierra Leone stands out for its positive results, measured by a "standard index of general economic welfare" based on fourteen key outcomes including average household principal components analysis asset score, average number of income sources per household, and average total income of households and an additional 19 outcomes (such as average agricultural income contingent on presence of farmers, or average days in school contingent on children). The studies from Aceh and Afghanistan find mixed economic welfare results. The Aceh evaluation finds a strong effect in reducing the share of households classified as poor. The other

indicators, such as the asset index score, the employment rate, and perceptions of improvement do not emerge as significant, although some of the individual assets in the index do increase significantly. In Afghanistan, the results are mixed in a different way: while there is no substantial effect on income, consumption or unemployment, both men and women perceived that their economic situation had improved over the past year.[13] The studies from DRC and Liberia do not find positive economic results.

In regards to governance outcomes, in relation to Afghanistan's NSP, only one key governance index is yet publicly available. In creating a "standard index of men's perceptions that various officials act for benefit of all villagers," Beath et al. present estimates for 9 different types of officials, and they all show positive effects of 5 percentage points. The study of Liberia's CDD/R programme emerges with a positive result falling just below traditional levels of significance from its standard index of democraticness, including such measures as "attitudes on individual right to question leaders, clientelism, role of all in decisions, importance of elections, community as primary decision makers for collective problems" (Fearon, Humphreys, and Weinstein 2011 46). In Sierra Leone, the authors report on a standard index of development committee functioning as highly positive and significant (15 outcomes, plus 43 continent outcomes), although there were no significant effects on the standard index of local authority (25 outcomes, plus 29 contingent outcomes) or standard index of access to information on local governance (based on 19 outcomes plus 21 contingent outcomes). The evaluations of CDD/R programmes in Aceh and DRC fail to find significant governance effects. In Aceh, outcomes are not indexed and only one outcome emerges at nearly statistically significant levels: the share reporting that local authorities would be effective in improving the community's situation is higher in treatment than comparison communities. In DRC, a number of indexes and measures related to both the RAPID behavioural measure and the survey do not show significant governance effects of the Tuungane programme.

In regards to social cohesion outcomes, only the study of the programme in Liberia finds measurable positive impacts. Fearon et al. find positive results across a range of standard indexes: the inclusion of ex-combatants, inclusion of migrants, trust in leaders, reduced tensions, and social capital. The studies from Sierra Leone, DRC, and Aceh all fail to find positive effects on social cohesion, measured in a variety of ways across the studies. In fact, the Aceh study comes out with one significant adverse effect: individuals in treatment communities are less accepting of ex-combatants than those in comparison communities. The authors of the Afghanistan study do not yet report on social cohesion outcomes.

The studies' own reports on intermediate outcomes are more complicated in that not every study framed their findings in terms of empowerment/voice and infrastructure/social infrastructure (or

---

[13] In the case of Afghanistan, we draw on Beath, Fontini and Enikolopov 2011, 2012a, 2012b, 2012c, 2012d. The final report was not yet public at the time of writing and we chose not to focus on the midterm results for best comparability with the other studies.

peace dividends). In relation to infrastructure/service delivery, only the study on Sierra Leone, using a standard index of local public services infrastructure (16 outcomes including such things as community having a functioning primary school, health unit, hut, water well, and dry floor, plus 34 contingent outcomes), found clearly positive results. The DRC study reported on an important number of classrooms and health clinics built or rehabilitated, but showed no improvement in respondents' access to services. The Aceh and Liberia studies did not find positive outcomes in relation to service delivery or infrastructure, measured by the presence of public goods like a school and health clinic in Aceh and measured by both presence and access to such public goods in Liberia. The currently available literature from Afghanistan did not report on this outcome.

In regards to empowerment/voice, the study in Afghanistan shows strong positive results reported through male and female standard index of acceptance of women's participation in village governance, a standard index of women's perception that local leaders are responsive to them, and other indexes related to the role of women in society. Only on the index related to men's acceptance of roles for women outside the home were positive results not found. The Sierra Leone study shows strong positive results on one of the two indexes the study authors use that relates to empowerment/voice – the index measuring participation in local governance, but not on that related to inclusion and participation in community decisions. The study in Liberia finds mixed results on voice/empowerment: strong results on indexes of efficacy and participation, but a weaker treatment effect, unable to meet traditional levels of significance on a standard index of women's rights. The DRC study finds no positive voice/empowerment effects. The study in Aceh does not find improvement in voice/empowerment, and indeed a negative finding on the share believing that villagers play the most important role in community decision-making.

A special consideration for CDD/R, and any development programme, is the possibility of adverse effects. In the study on DRC, responses to a question on household income earned over the past two weeks represented a statistically significant decline.[14] In Aceh, as noted above, individuals in treatment communities are less accepting of ex-combatants than those in comparison communities. The share believing that villagers play the most important role in community decision-making was also lower in treatment communities (see technical appendix). Just as a multiplicity of outcomes may lead researchers to incorrectly report significant results in a desirable direction, the same may be true of an undesirable direction (see Humphreys, Sanchez de la Sierra, and van der Windt 2012, 75). In Sierra Leone, while there are no adverse effects reported on an entire index, sets of outcome variables do include some negative effects. For instance, of the 59 outcome variables related to collective action and contributions to local public

---

[14] A couple of other findings also emerge as adverse and nearly statistically significant such as access to roads. The authors explain that, "in the context of the multiple measures given here this single adverse finding is plausibly due to chance" (Humphreys, Sanchez de la Sierra and van der Windt 2012, 75).

goods, 7 treatment effects meet conventional levels of statistical significance: 5 positive and 2 negative, which the authors interpret as "a zero effect when considered as a group" (Casey, Glennerster, and Miguel 2011b, 37). The section on "is it evaluation strategy?" below, elaborates on the issue of adverse effects.

The overall reading is that while CDD/R has nearly no adverse effects, it is better at generating more tangible economic outcomes than it is at generating "softer" social changes, although not all programmes were effective in producing economic outcomes. This interpretation speaks to the difficulty of effecting social and political change. It is broadly consistent with the findings of a recent review of CDD interventions by the World Bank, including non-conflict affected contexts (Wong 2012).

An alternate way to think about these findings is to do more to separate out proximate impacts associated with the project (or what the ToC characterizes as intermediate outcomes) from more distant outcomes separate from the project (the ToC's final outcomes). Here, as illustrated in Figure 8, a slightly different and more consistent picture emerges. Upon closer examination, many of the economic measures from Sierra Leone and Aceh, where positive economic results were found, are quite close to process or proximate output measures, with some measures relating directly to the programme. Under this interpretation, many, but not all, of the positive economic results from Sierra Leone and Aceh can be moved up to the social infrastructure/peace dividend category, or even into outputs. In Sierra Leone, for instance, "community has a bank account", "attended skills training" and "any petty traders in village" are included as economic welfare indicators (Casey, Glennerster, and Miguel 2011a, 37). In Aceh, since communities had the opportunity to disburse grants to individuals and groups, which they did primarily for economic purposes, rather than spend on public goods, specific outcomes including ownership of engines and motorcycles and an increase in farming productive land can be interpreted as quick wins, although farming is a more distant outcome than purchasing engines. As Barron et al. write, "there is, however, no direct evidence of welfare gains in other areas such as school attendance, health and employment levels. Unsurprisingly, given the small proportion of project funds used for public goods, the programme is not associated with changes in the level of community infrastructure" (Barron et al. 2009, 3).

The picture presented by framing the results this way suggests that CDD/R is better at producing outcomes directly associated with the project. Outcomes that are more immediate and more directly connected to the intervention, such as social infrastructure or short-run economic effects, are more achievable. CDD/R as currently conceptualised and implemented is not good at inducing changes in routine social life. In thinking about the theory of change, this means that the learning by doing mechanism does not appear to be transpiring.

Figure 8: A Reinterpretation of Findings on Key Outcomes from Five Impact Evaluations

Outcomes		Afghanistan	DRC	Aceh (Indonesia)	Liberia	Sierra Leone
Proximate	**Peace dividend (social infrastructure, service delivery, short-run economic)**		🟨	🟩	🟨	🟩
	Empowerment/voice	🟩	🟥	🟥	🟨	🟨
Final outcomes (indirect)	**Improved economic welfare**	🟨	🟥	🟨	🟥	🟨
	Improved governance	🟨	🟥	🟥	🟨	🟥
	Improved social cohesion		🟥	🟥	🟩	🟥

Further strengthening this interpretation, the Liberia research team qualifies their positive "green" finding on social cohesion since that finding is entirely in places where mixed gender groups played their public goods game. In this interpretation, discussed further below, the demonstration effects for routine life are limited since results did not extend to women only groups. Similarly, in Afghanistan, at least vis-à-vis the results currently available related to women's involvement, the strongest results are those more directly linked to the intervention, such as men's and women's perceptions of women's role in governance, not those more distant to the NSP intervention, such as men's perceptions of women's roles outside the home.

The evaluations do relatively little to unpack why the programme works or does not in accordance with this or other theories of change. As Humphreys writes, a "counterfactual (RCT) approach can estimate effects whether or not you know why effects would arise"(Humphreys 2012). If the interventions were producing strong positive results, some interviewees would be less concerned about the why questions. Nonetheless, since results are mixed the why question takes on particular importance. The Liberia study goes the furthest in this regard, reasoning through different causal pathways to explain the fact that positive social cohesion effects

emerged only in the mixed-gender, rather than women only, public goods games. They hypothesize that "the main effect of the CDR programme may have been to increase the number of community members with leadership and organizing experience relevant to solving a similar community-wide collective action problem. There is also some evidence that CDR-treated communities used more democratic processes to select community representatives and projects" (Fearon, Humphreys, and Weinstein 2011 3).

Overall, this synthesis of CDD/R projects does not lead to the kind of impacts one would need for CDD/R, as currently designed, implemented, and evaluated, to be considered a "proven impact" intervention.[15] While the interventions have been tested using scientifically rigorous methods, the theory of change is inadequately clear, the results are mixed, and we know little about cost-effectiveness or effectiveness of CDD in delivering goods, or improving governance, social cohesion and welfare, compared to other development alternatives.

We had a special interest in learning from CDD/R in conflict-affected contexts and wanted to know what rigorous impact evaluations could tell us. Would the review findings have been different if the inclusion criteria for this review were wider? For instance, what do the studies from Philippines[16] and Indonesia, with some parts of each country plagued by conflict, conclude? The evaluation of the Philippines' KALAHI-CIDSS shows a positive impact on household consumption and improvements in basic service delivery. The study also notes an increase in barangay (village)-level assemblies and greater knowledge about barangay income and expenses, as well as a positive impact on group membership and trust, but a negative impact on collective action (Edillon, Piza, and Santos 2011). A study of social capital specifically finds increased generalized trust levels, but a negative impact on collective action and group membership (Labonne and Chase 2008). In Indonesia, results from an evaluation of phase two of KDP (2003-7) suggest that the programme had a significant impact on consumption in poor households, and on unemployment rates, as well as a positive impact on access to health, but not on school enrolment rates (Voss 2008). The evaluation of PNPM Generasi (the next phase of KDP, 2007-2010), a programme targeting explicitly health and education, found positive effects on all of its indicators, with the most significant impacts on those with low baseline measures (Olken, Onishi, and Wong 2011). These findings are thus relatively consistent with the interpretations above, including CDD being better at service delivery and direct outcomes, and producing more mixed results, where it has been measured, on social outcomes separate from the project.

The other two recent reviews relating to CDD/R are also generally consistent with the interpretation that emerges from the studies included here. In Mansuri and Rao's (2012) book on

---

[15] See, for example, http://www.poverty-action.org/provenimpact/testedandproven.

[16] Please note that I am a consultant working with *Innovations for Poverty Action* as part of a team to conduct the evaluation of the second round of the KALAHI-CIDSS CDD programme in the Philippines (2011-2016). No results are yet available.

"induced" participation, drawing on almost 500 studies of participatory development and dencentralization, the authors note that "the process is, arguably, still driven more by ideology and optimism than by systematic analysis, either theoretical or empirical" (2012, 3) While Mansuri and Rao identify some things at which participation is good (such as improving resource sustainability and the quality of infrastructure, (2012, 6)), they identify a number of ways in which participation fails to live up to its promise (such as building long-lasting cohesion, (2012, 9)). They also find significant heterogeneity in outcomes, even across small sets of studies. Wong (2012) has a more narrow focus on CDD, examining only World Bank-funded projects evaluated to rigorous impact evaluation standards. She finds that across the programmes she studies, there is generally strong evidence for improved access to, and use of, services such as water and education. Wong notes that these are the principal stated aims of the majority of studies she includes, a notable difference from the programmes studied here. She finds that many studies do not expect broader socio-economic welfare effects, but among those that do and test for them, the economic wellbeing effects of CDD are fairly well founded. In contrast, the evidence for social cohesion or governance effects, again, not frequently tested, is more mixed or lacking. Wong adds, in reference to programming in conflict-affected states including the GoBifo, Aceh, and NSP programmes included here, as well as the programmes in Indonesia and Philippines here excluded, that "not surprisingly, there is no impact on macro levels of violence except in the case of insurgency groups in the Philippines"[17] (Wong 2012, vi). Findings from both reviews are thus quite consistent with the findings here, although the results may be even more circumscribed in conflict-affected contexts.

Discussion: Explaining the results

What factors contribute to explaining these rather disappointing results? There are many likely explanations, with overlaps among them. The conclusion of the authors of the Sierra Leone study warrant repeating: "...there is much that we, as outsiders, do not yet know about how to do this [CDD/R] effectively" (Casey, Glennerster, and Miguel 2011b, 46). At this point, it is not possible to figure out which of the following explanations is most responsible, and the answer may vary from case to case; it would seem that the answer is to some extent "yes" to all of the following questions.

Is it the goals?

CDD/R is marked by *panacea-type thinking*. CDD/R as an approach is a response to perceived failures of top-down, donor-driven, development and peacebuilding projects in alleviating poverty (International Development Association 2009). Many interviewees believed that CDD/R

---

[17] One article focusing on the Philippines' KALAHI programme uses a rent-seeking model of conflict and a geo-referenced dataset on the occurrence of conflict between 2003 and 2006 matched to household surveys and budget data from municipalities across the Philippines. The authors find that KALAHI-CIDSS leads to a decline in events related to the Moro Islamic Liberation Front and an increase in events related to the New People's Army, rebel groups with differing ideologies. The authors hypothesize that while the former may feel included by KALAHI, the latter may be an effort to disallow a peace dividend (Arcand, Bah, and Labonne 2010).

was good for something. Nonetheless, the grand, interconnected, goals, paired with the seeming simplicity of the pathways, and the commonly quite short CDD/R programming may be setting community-driven development and reconstruction initiatives up for failure. Some interviewees suggested that CDD/R is suffering from "mission creep" from earlier conceptualisations as a service delivery mechanism to embrace broader social and economic goals. Indeed, the notion that what often remains one "project" can change this many deep-seated variables is unlikely. If CDD/R is theorized and undertaken as a panacea, there is a danger that they may all fail due to unrealistic expectations (Mansuri and Rao 2012, 12; Haider 2009 14-5).

Is it the theory of change?

Interviewees often noted that the lack of one clear, driving ToC is a potential strength of the CDD/R approach in that, just as the programme can be adapted in many different ways, the ToC can be adapted to match. Generally, though, that the theory (or theories) of change remain underspecified, is a weakness in the CDD/R literature and programming. Interviewees described the generic theory of change as "lofty", "unrealistic", "inherently flawed" and even "ridiculous". In this way, the weak theory driving programming may help explain the rather disappointing CDD/R results.

In particular, contrary to the **simplistic** ToC presented above, the processes that CDD/R programming attempts to put into action are unlikely to work in a straightforward fashion. Rather, complexity theory can inform our thinking. The trajectories toward the outcomes of interest are likely to be characterized by multiple possible roads to the same outcomes, tipping points (or critical thresholds), and non-linearity, meaning that the strength of an output or outcome does not vary consistently with the strength of the input. As Mansuri and Rao similarly argue, "effective civic engagement does not develop within a predictable trajectory. It is instead likely to proceed along a punctuated equilibrium where long periods of seeming quietude are followed by intense, and often turbulent, change" (2012, 12). Practically, it is difficult to craft something like the log frames common to development practitioner work when there is dialogue and interaction between different trajectories and objectives.

The idea that quite brief exposure to participatory, democratic, governance practices will change attitudes and behaviour is overly ambitious. As the authors of the Tuungane evaluation put it, this "basic principle behind CDR…runs largely counter to classic accounts of the determinants of social behavior that emphasize structural and **slow moving features**" (Humphreys, Sanchez de la Sierra, and van der Windt 2012, 19). The authors of the Liberia study acknowledge "a large academic literature that casts doubt on the ability of foreign aid to have much real impact on deeply rooted patterns of social cooperation, conflict, and local governance structures" (Fearon, Humphreys, and Weinstein 2008 39). Many interviewees were very critical of the idea the short-term intervention could translate into long-term real-world changes to the community at large.

This theory of change is based on relatively **static inputs**. In contrast, CDD/R programming should, by definition, be relativistic and evolutionary. Contextual moderating factors will affect the pathways through which change can happen and CDD/R design should adapt as it matures.

One purpose of theories of change as it relates to evaluation is to articulate what one will test and measure. The **outcomes are broad and some are quite nebulous**, and all can be measured in a host of different ways, as discussed above.

Trajectories of change are likely to depend on the specifics of the programme in question.

For example, programmatic features, such as the intensity of facilitation and the number of rounds of funding, may more strongly or weakly lend themselves to learning by doing. The openness of the menu of programme choices may also affect the trajectories of change: a list that is too restrictive may produce artificial rather than genuine empowerment (Haider 2009 9).

The theory of change and evaluations need to do more to match project type to theories and outcomes. Figure 9 begins to engage in this exercise. If communities choose to use their block grants on educational infrastructure, such as rehabilitating schools, for example, it does not logically follow that income or asset indicators will increase in the short term. In Liberia, many communities spent their funds on recreation rooms and guest houses for people to stay when they travelled through the country. Income generation projects were explicitly excluded from the menu (IRC 2008, 11). It does not follow, therefore, that economic wellbeing would likely be a primary outcome of these projects. But, we may expect gains in some sense of community. In Aceh and Sierra Leone, the bulk of funds was spent on sub-projects dedicated to skills training and income generating initiatives, so we would expect more direct and immediate economic outcomes. In Aceh, given the focus on private goods, there was little opportunity for collective implementation: "Many of the goals of CDD may depend upon processes that are brought into play conditional on particular types of activities (joint selection of projects, community oversight of implementation, etc.). Yet certain kinds of activities that communities might choose are less likely to encourage interaction, limiting some of the gains that CDD projects might purport to have." (Barron et al. 2009 iv)

Figure 9: Matching inputs and outputs

Case	Two top subproject categories	% funds spent	Significant project outcomes
DRC	Education	45%	None
	Water and sanitation	18%	
Aceh (Indonesia)	Economic (mostly individual direct transfers for cattle, agriculture, etc.)	83%	Proximate economic
	Infrastructure	10%	
Liberia	Health water and sanitation	46%	Social cohesion
	Social projects	32%	
Sierra Leone	Economic (Agriculture 26%; Income generation/training 17%; Livestock/fishing 14%)	57%	Proximate economic
	Community centres/sports	14%	

Participation is at the core of CDD/R. The idealised thinking about participation embodied in the theory of change is sometimes deemed the "hippy model" of participation "as it appears to suggest that all community members enter the decision making sphere on an equal footing, and can agree a common interest without intra-community conflicts" (Vajja and White 2008, 1148). Shifting existing power arrangements is similarly presented as a relatively straight-forward, simple, easy, thing to do. But, there are costs to participation, including "effectively shift[ing] some of the burden of service provision onto beneficiaries" (Casey, Glennerster, and Miguel 2011b, 9; Cooke and Kothari 2001). In an internal process evaluation of the Liberia CDD/R programme, for instance, authors note that they have "difficulty envisioning collective action toward a public good when people have little free time or resources to spare" (IRC 2008, 10). In addition, challenging the status quo is a conflictual exercise: voicing different views may exacerbate conflict and participation may be threatening to elites (Haider 2009 7; Cliffe, Guggenheim, and Kostner 2003). The theory generally underplays the radicalness of social change (Mansuri and Rao 2012, 98-9).

There are **tensions with the timelines** for CDD/R. While the theory of change presents proximate, more short-term goals as leading to long-term goals, there may be a tension between these two sets. As Slaymaker et al. write, "a common challenge in post-conflict settings is striking an appropriate balance between the need to rebuild institutions quickly and the desire to reform them to ensure longer-term sustainability in service provision" (Slaymaker, Christiansen, and Hemming 2005, 4; see also Strand et al. 2003, 7; Mansuri and Rao 2012, 292; IRC 2008, 15). The authors of the Aceh study note possible trade-offs of adverse short term effects to bring about positive long term effects: "If programmes like BRA-KDP increase the counter-veiling power of civilian community members, it may create an important check and balance to GAM power… Conceptualized as such, an increase in minor tensions between former combatants and civilians may be positive for post-conflict stability, as long as such tensions do not escalate into

more serious conflict as they rarely have to date" (Barron et al. 2009 65). On timing more generally, this theory of change does not suggest a timeline over which effects are expected to generate.

There are **other potential pathways from inputs to outcomes** that merit consideration. In a report on IRC's involvement in the NSP in Afghanistan, Maynard suggests that hiring facilitators from each local area increased jobs in those provinces, and thus income and savings. "The fact that NSP staff live in the same district in which they will work implies the skills they gain through training and experience will remain in the area….In the medium term, however, IRC's approach [to hiring] appears to support project sustainability and increased self-reliance" (Maynard 2007 11).

Is it implementation?

If the projects were poorly implemented, we would not expect to see strong results. Assessing the quality of project implementation with existing evidence is tremendously difficult. Some of the evaluations measure basic implementation suggesting that in DRC, for instance, the "projects were implemented according to plan" (Humphreys, Sanchez de la Sierra, and van der Windt 2012, 15) and that in Sierra Leone "the programme did what it said it would" (Casey, Glennerster, and Miguel 2011a, 1). In the case of Liberia, an internal process evaluation notes a number of implementation difficulties (IRC 2008, 2-4). Overall, we often have little detail of the quality of implementation. According to interviewees, much implementation monitoring entails a simply binary yes/no, rather than a quality scale.

Of projects that did evaluate implementation – already going a good way from the more black box projects in which we know very little about implementation quality – there are still a number of important implementation quality issues left out. The Tuungane project, for example, included timing of project start, delays in project identification, delays in project implementation, type of project selected, existence of complaints about disbursements and project suspension due to conflict as indicators of quality of implementation (Humphreys, Sanchez de la Sierra and van der Windt 2011). But, what is the quality of the subprojects? It makes sense that stronger outcomes will follow from better public goods. The engineering matters but none of the evaluations here systematically measured quality of projects. The Sierra Leone study reports on perceptions of subproject quality (Casey, Glennerster, and Miguel 2011a, 32). In interviews, several of the study authors and programme implementers noted that project quality was quite poor. There are particular concerns regarding places where the programme required community labour for technically difficult projects, such as roads and irrigation, with little quality control. Other important factors include the quality of participation and the quality of facilitation. Moreover, what makes for good quality on these indicators?

Quality of implementation can be rather subjective. In Sierra Leone, for instance, disbursement of the block grants were sometimes so delayed that community members called the project

"GoBien" (backward march) instead of GoBifo (forward march) (Sulley et al. 2010, 60) and in other programmes as well, many projects were delayed. Many of the CDD/R programmes were plagued by seasonal challenges, difficulty accessing remote villages and lack of qualified contractors to execute projects.

Quality of implementation may also be influenced by the fact that there is an evaluation. In Afghanistan, the evaluation team conducted monitoring of elections, sub-project selection, and grant disbursement, but believe they may have produced better implementation by the very inclusion of communities in the evaluation.

It is output strength?

If outputs are weak, we may not expect strong outcomes. As noted in the theory of change discussion above, the outputs are ambitious – things like participation, changing local power structures, community determining priorities, and projects reflecting people's priorities. These outputs are also difficult to measure and interpret. While evaluations often provide statistics such as attendance rates at community meetings it is tremendously difficult to make sense of findings. The DRC evaluation reports that 30 % of the total population, 23% of women and 36% of men attended at least one meeting. The Aceh evaluation reports that 20% attended meetings (Barron et al. 2009 28). In Sierra Leone, 64% of respondents said they attended the meeting where the Village Facilitator was chosen and 41% said they attended a meeting to draft the Village Development Plan (Casey, Glennerster, and Miguel 2011b, 32). Do these represent "good" rates of participation? What about quality?[18] In some cases, it is not clear whether participation rates should be read as inputs or ouputs insofar as minimum participation rates are mandated in some cases, such as requiring 50% of eligible voters to vote in a referendum in order for it to be valid in Afghanistan's NSP.

Other key outputs in the theory of change include the community determining investment priorities and projects reflecting people's priorities. In Afghanistan, elites had significant sway over project choice when choices were made through at-large elections and consultative meetings; this was less the case when choices were made via referendum (Beath, Christia, and Enikolopov 2011). In Liberia, there was no significant difference between the average share who reported selected projects were among "most important" for the village in treatment and control communities (Fearon, Humphreys and Weinstein 2008, 29). Similarly, the "share agreeing that 'an ordinary person can influence decisions about community affairs', the "share agreeing that they are 'influential' in decision-making" and the "share claiming to be satisfied overall 'with the way that decisions that affect all community members are made'" was not significantly different in communities exposed to CDD/R than in others (ibid. 17). In Aceh Timur, the central highlands and west coast Aceh, non-elites and disadvantaged groups, including conflict victims, often did

---

[18] Note that several evaluations measure quality of participation in their post-intervention behavioral measures. These are not measures of intervention output strength.

not attend the village meetings because they were not invited. In the same areas, evidence suggests that village elites and ex-combatants dominated the decision-making process. However, in Bireuen, key target groups were able to participate frequently (Morel, Watanabe, and Wrobel 2009 27). 65% of villagers report that their preferred project was approved – likely because the approval process was competitive – but this figure is 6 percentage points lower for conflict-affected individuals and 10 percentage points lower for the most-conflict affected (Barron et al. 2009 27-8). In DRC, nearly 70% of all projects were matched to villagers' preferences (Humphreys, Sanchez de la Sierra, and van der Windt 2012, 24). In Sierra Leone, respondents reported that roughly 38% of these decisions were made by the chief and other leaders, with for 43% of decisions made by "everyone", 1% made by an outsider and 18% unknown (Casey, Glennerster, and Miguel 2011b, 32). Here too, it is difficult to judge the strength of these results.

Is it programme design?

A number of different programmatic factors may affect outcomes, and this list is, of course, not exhaustive. First, the DRC, Aceh, and Afghanistan studies all suggest that their respective lack of results, on all or certain outcomes, may be explained by the **relatively short length of the project**. They hypothesize that a longer programme could have more impact; as the authors of the Aceh study write, "one reason may be that the programme only ran for one year, limiting the extent to which such gains, which tend to build over multiple programme cycles, could eventuate" (Barron et al. 2009, 62). This line of argumentation is consistent with broader literature about slow changing socio-political processes. The KDP project in Indonesia (separate from BRA-KDP included here), for instance, found that impacts on social relations were stronger in the third and fourth year of the programme than they were earlier on (Barron 2010, 21). A survey of World Bank staff asked "in your experience, what would be the average number of years needed for project support of community groups initially formed under the process to reach a level of sustainability of community processes requiring very limited outside support (such as simply a supporting/maintenance visit once a year)?" 52% of respondents suggested that the timeframe was six to ten years, with nearly 24% suggesting it was one to five years and the rest suggesting eleven years or more (Kumar 2005, 101). Another World Bank publication suggests that CDD/R requires more time in conflict-affected contexts than elsewhere since the infrastructural, institutional, and social needs are great (World Bank 2006, 22). At the same time, interviewees cautioned that the timeline needs to be paired with intensity. Stretching the same amount and quality of facilitation or amount of funding or number of funding cycles over a longer period of time is not what they meant by increased timeline. Given different contexts and goals, it is likely that programme length will need to differ. That "patience is a virtue" in CDD/R (Mansuri and Rao 2012, 306) comes into tension with time pressure to disburse and show results inherent in the development industry.

Second, the studies and interviewees hypothesized that the **size of block grants** is too small to effect meaningful change. The DRC study notes that "for economic outcomes the low per capita investments… may plausibly account for the results" (Humphreys, Sanchez de la Sierra, and van

der Windt 2012, 8). In contrast to the huge sums quoted as spent on CDD initiatives, Figure 10 shows the size of the block grant per capita, which, among the cases studied here, tops out at about $20.40 per capita in Afghanistan over the entire duration of the project. This can be compared, for illustration purposes, with the Millennium Villages project which aims to spend $120 per capita per year (Earth Institute 2008). In some cases, village priorities had to be redirected since the funds would not be enough to complete their preferred project, despite community determining priorities and investment reflecting people's priorities being a key part of the ToC. There may also be a threshold (or tipping point) at which effects transpire. A few interviewees felt that the size of the grants was appropriate, and that if they were larger, the village would not be the appropriate unit at which to disburse them. Many suggested that multiple grants for repeated exposure to the project and better learning by doing would produce better results than one or two rounds. It is also believed, but not tested in these evaluations, that requiring community contributions are important for bringing about results.

Figure 10: Investment per capita[19]
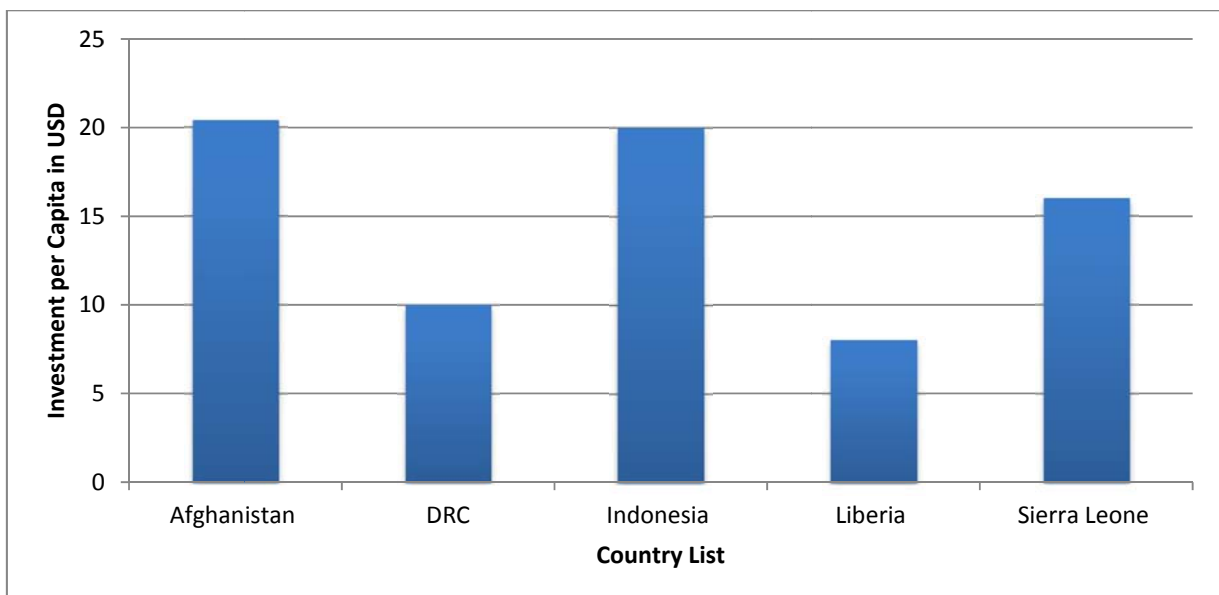
Third, and related, the **limited reach of selected projects** may explain the lack of findings on certain outcomes. In the DRC study for instance, the project was implemented over a huge territory. With upwards of 1,700 classrooms and 150 clinics built or rehabilitated across eastern DRC, even if they service entire villages, the clinics will serve about 10% of the population and the classrooms will directly affect about 900,000 students, at the most 5% of the population

---

[19] Calculations are over the course of the entire project, not per year. The $20.40 for Afghanistan is calculated by taking $200/household and dividing this by an average of 9.8 people in the household (Beath, Christia and Enikolopov 2011b, 31). The figure for DRC is calculated at $10 per capita ($1 per year x 2 for lowest level projects + $4 per year x 2 for higher level projects) and Aceh is taken as $20 per capita (Humphreys, Sanchez de la Sierra and van der Windt 2012, 18). Liberia is reported to be $8 and Sierra Leone is reported as $16 (Casey, Glennerster and Miguel 2011b, 14).

(Humphreys, Sanchez de la Sierra, and van der Windt 2012, 18). The reach of the project may have to be larger or stronger to effect change.

Fourth, **the focus on social infrastructure may limit outcomes**. Despite much of the literature referring to infrastructure and service delivery, most of what happens in CDD/R programming relates to infrastructure. This is important, but inadequate. If a classroom is built, but there is no teacher or didactic material, its impact will be limited. As one interviewee added, if infrastructure for health or education is provided without a strong nation-wide programme for subsidizing costs, those services become extremely expensive for people and CDD/R may in this way actually exacerbate rather than decrease inequalities.

Fifth, **menu restrictions for projects** may affect outcomes. CDD/R implementing agencies typically limit, at least in some ways, the menu of subprojects from which communities may choose. This specification introduces significant tensions. First, if the community determining investment priorities – thereby inducing accountability, incentives to economize, and collective ownership – is indeed a key part of a theory of change, as suggested in the generic version above, having the community choose how to spend their block grant is indeed important. An even partially restrictive menu may block this pathway. For example, in all cases, except Aceh, programme design required that funds be spent on public rather than private goods. On the other hand, by allowing people a larger choice of projects, as in Aceh, they may select themselves out of some of the key social aspects of the programme. Other restrictions, while meant to bolster one outcome, may at the same time hinder the achievement of others. For example, in Liberia, it was a rule that the funds not be spent on income-generation projects, such as capital equipment for rice mills. This may result in more projects from which the entire community can benefit, prioritizing social cohesion, but it may limit economic outcomes.

Sixth, the **quality and intensity of facilitation matters**. This includes both planned activities and variation introduced by individual facilitators. In terms of planned activities, some interviewees suggested that programmes may need to spend much more time and money on facilitation, making this component of the programme more intensive. Facilitation varied significantly across the projects under study here from a minimum requirement of four meetings in Aceh to roughly six months of direct facilitation over four years in Sierra Leone. That the type of facilitation may need an overhaul, depending on the goals, also arose in several interviews. If, for instance, social cohesion were paramount, perhaps projects should focus more on social interaction. Finally, the skills, approach, and commitment of individual facilitators varies tremendously. Individual facilitators are central to the CDD model, but they are little supported and sometimes do not even attend training, such as in Aceh, where it was too expensive to travel for it (Morel, Watanabe, and Wrobel 2009 51-2). In conflict-affected contexts, the positionality of facilitators matters too such that they are not seen to be on one side of the conflict (Haider 2009 17; Cliffe, Guggenheim, and Kostner 2003, 13); in Afghanistan, the most effective

facilitators have been from the same groups as the communities in which they work (Barron 2010, 24).

There is good reason to **question the effectiveness of gender quotas**, at least as currently implemented in programming. The Afghanistan, DRC, and Liberia programmes included specific quotas for women on councils.[20] In Sierra Leone, facilitators "actively encouraged women and youth to participate" and one of three co-signatories for community bank account had to be a woman and in Aceh, facilitators convened a "women's meeting". While in each of these cases, this ensures women a "seat at the table", there were limited or no positive attitudinal spillovers regarding the roles or responsibilities of women in society (Beath, Christia, and Enikolopov 2012b 3; Humphreys, Sanchez de la Sierra, and van der Windt 2012, 8; Casey, Glennerster, and Miguel 2011b, 40). Interviews suggested some potential explanations. In DRC, it was a requirement that the CDC have two co-presidents, two co-secretaries and four ordinary members; in effect, this may have resulted in a "full position" for men, with a "vice" or "co" position for women rather than giving women full title, rights and responsibilities. There is also the risk of partiality or bias (Haider 2009 8). Others suggested that dividing the pie runs counter to the theory of change which focuses on participation by all and cooperation. In Afghanistan, where 15% of funds were specifically dedicated to a women's project, interviews suggest that this resulted in men (50%) controlling 85% of the funds, and women (50%) controlling 15% of funding, implying that this ear-marking disempowered rather than empowered women. Yet, more than one of the studies cites positive evidence for quotas (from non-CDD/R programming) in India (Beath, Christia, and Enikolopov 2012b; Casey, Glennerster, and Miguel 2011b), acknowledging however that that differences come only after two rounds of quotas (Beath, Christia, and Enikolopov 2012b, 24), relating this issue to the length and intensity of CDR programming and to institutional context, discussed below. Broader cultural context and the pre-existing role of women in society also matters, of course.

Seventh, explanations may lie in the **way that communities are conceptualised**. The concept of community is at the core of the CDD/R approach. The World Bank defines a community as "a group of people who share broad development goals. Their social behavior and relationships are governed by social norms that are expected to provide solidarity" (Kumar 2005, 55). That communities do, or have the potential to, share goals and relationships or a "civic mentality" (Haider 2009 19) are important assumptions underlying CDD/R programming. The CDD/R programmes reviewed here are all based on geographic conceptions of community. Participants may not define their communities geographically.

Moreover, how communities are defined and constructed differs from programme to programme. Some implementing agencies themselves define geographic communities for the purposes of their project. For a variety of reasons related to funder requirements of the number of people to

---

[20] In Liberia, this was specifically set up as an experimental intervention.

be served, logistical capacity, and distribution of villages and people across the selected geographic area (Fearon, Humphreys, and Weinstein 2011 7), implementers may merge or divide naturally existing groupings for the purposes of the project. In Liberia, for instance, villages were grouped into equally sized "communities" based on proximity and existing ties to make 2,000 to 3,000 person communities (Fearon, Humphreys, and Weinstein 2008 2). In DRC, researchers report that IRC wanted to implement in 280 communities of 6,000 people each, meaning that they had to add and divide communities. Nonetheless, these regroupings did not result in revised assumptions about the community or its potential. If communities are themselves fabricated for the purpose of the intervention, the institutions created are unlikely to have any meaning outside of the programme. Additionally, that the corollary of an inclusive community is that "those who do not belong to that community are 'excluded'" is rarely acknowledged (Kumar 2005, 55).

The definition of community also brings up a tension as it relates to conflict-affected contexts. To some extent, the definition of community noted above "neglects community members' differences and power relationships, the conflicts, and the diversity of interests that determine day-to-day behavior and that have an impact on the effectiveness of participatory approaches" (Kumar 2005, 55). At the same time, and somewhat paradoxically, in conflict-affected contexts, it often comes across as if the different parties to conflict all live in the same community. This presumption appears to underlie some of the social cohesion and peacebuilding goals of CDD/R. However, in places like Afghanistan, Congo, and Liberia, in contrast to somewhere like Rwanda where, in many communities Hutu and Tutsi do indeed live side by side, the key lines between fighting factions do not fall at the village level. There are, of course, divisions within communities, and local level conflict is important in these contexts (i.e. see Autesserre 2010 for the DRC), but these may not reflect the meta-conflict cleavage. Programming needs to understand these divisions, and base planning on definitions of the community, and who is in it.

Related questions arise around the possibility of determining an ideal community size and administrative level for CDD/R programming. In Sierra Leone, staff pondered whether smaller communities were better able to make use of the CDD/R model than larger ones, although the study authors did not find differential treatment effects in their study (Casey, Glennerster, and Miguel 2011a, 23-4). Figure 11 illustrates the size of communities in each of the cases considered here. Some interviewees also wondered about factors such as geographic size of the community and proximity to economic centres. There is also the challenge of ensuring the CDD/R is targeted at the appropriate unit for collective decision making. The GoBifo project, for instance, involved projects at the village and higher ward level, with the evaluation focusing on village level outcomes (Casey, Glennerster, and Miguel 2011b, 43-4). The Tuungane programme also involved committees and grants at two different levels, with the evaluation focusing only on the lowest level. The authors of the Tuungane evaluation pondered the "…possibility…that the programme is pitched at the wrong level to effect change in governance structures and social cohesion; Tuungane has focused on the most local levels which may not display the same

problems of cohesion and weak governance that are so visible in Congo at the macro level" (Humphreys, Sanchez de la Sierra, and van der Windt 2012, 8).

Figure 11: Approximate number of individuals per community[21]
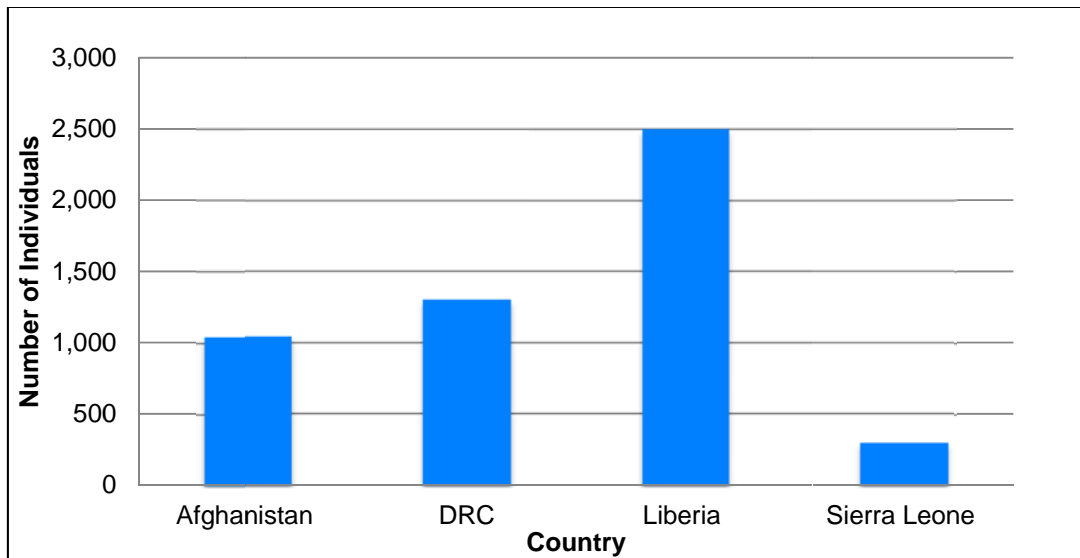
Eighth, whether or not the **community institutions are new or build upon existing institutions** likely matters. In Sierra Leone, for instance, GoBifo worked with the state and chieftancy system and did not make an aggressive challenge to the status quo. In contrast, in Liberia, "programmes have, as an explicit purpose, the goal of challenging existing structures of authority by putting power and resources in the hands of community members, especially previously disadvantaged or marginalized groups" (Fearon, Humphreys, and Weinstein 2008 1). In addition, in terms of outcomes, if the CDD/R programme mimics already existing institutions, the logic of its ability to induce demonstration effects is undermined. In DRC, for example, the likelihood of using elections was roughly the same in CDD and non-CDD/R communities, showing that elections were not newly introduced via CDD/R programmes (Humphreys, Sanchez de la Sierra, and van der Windt 2012, 7). Also, learning by doing for use in the future only makes sense if there will be a chance to use the learnings sometime in future. If CDD/R appears to be a single "project", individuals may try to get as much money as they can while the project lasts rather than challenge elites and embedded decision-making processes.

---

[21] For Afghanistan, the calculation of 1,044 is based on an average of 106.5 households per community (the average of 103 in control communities and 110 in treatment communities), multiplied by 9.8 as an average number of inhabitants per household (Beath, Christia and Enikolopov 2012b, 31). In, DRC the estimate is 1,300 (Humphreys et al. 2012, 12-3). In Liberia, we took the average of 2,500 from 2,000-3,000 individuals per community (Fearon, Humphreys, and Weinstein 2008 2). In Sierra Leone, the number was approximately 300 (Casey, Glennerster and Miguel 2011b, 14).

Related, **how the community-level institutions relate to the state** was perhaps the most frequently raised programmatic factor in interviews. One can contrast programmes in Liberia and DRC, conceived and run by NGOs, to the NSP, GoBifo and BRA-KDP, which were institutionally lodged within government departments. As Cliffe et al. wrote a decade ago, "two contradictory risks may occur with the role of government – that it is either too close or too distant from the CDR process" (Cliffe, Guggenheim, and Kostner 2003, 20). If CDD/R programming relies entirely on existing institutions, they risk becoming inadequately transformative and simply replicating existing power structures, which is counter to the ToC. If, however, they create entirely new institutions, they may induce conflict with elites; be irrelevant further down the road; or they may undermine the legitimacy of government structures which is important to build in conflict-affected contexts. There is usually a greater focus on using CDD/R projects to establish new semi-permanent bodies in places where there are not already government institutional bodies in place at the local level: The Sierra Leone study posits that its lack of findings on social cohesion and governance outcomes, in contrast to the CDR programme in neighbouring Liberia may be due to the fact that "the Liberia programme 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 programme did. Attempts to create new institutions and norms where formal structures have broken down may encounter less resistance than efforts to persuade existing authorities to adopt new practices" (Casey, Glennerster, and Miguel 2011b, 10).

In short, the specifics – even seemingly small specifics – of programme design matter. In the Afghanistan evaluation, for example, the evaluation team studied different treatment arms for elections (candidates by geographic cluster vs. election at large) and subproject selection procedures (secret-ballot referendum vs. consultative meeting). The authors compared types of projects selected and prioritized in different arms with ex-ante preferences of different groups of villagers to look at elite capture and indeed found different results with these differences in programme designs (Beath, Christia, and Enikolopov 2011). The conclusion, below, suggests future research related to different treatment arms for programme design.

Is it evaluation strategy?

As the authors of the DRC study recognize, "whenever research produces null results there is a natural question whether the nonfindings reflect a reality of weak effects or reflect shortcomings in research design or implementation" (Humphreys, Sanchez de la Sierra, and van der Windt 2012, 11). It is certainly a question worth asking. In terms of detecting impacts, the focus of experimental designs, the evaluations included here are strong. Nonetheless, a number of questions about evaluation strategy remain, the answers to which could influence evaluation moving forward.

Are the measures and instruments the best ones? Researchers typically spent substantial time piloting and testing the appropriateness of different measures and instruments. But, choosing the

right measures is a challenge, all measurement strategies are subject to some kinds of bias, and different instruments can lead to different results. The use of multiple and diverse instruments helps lessen the chance that any specific measure is weak.

The behavioural measures used in the evaluations studied here provide an illustration of the challenges related to measures and instruments (see again Figure 6). The Liberia study is the first randomised evaluation of CDD using behavioural games. The Liberia evaluation found positive effects on social cohesion exclusively in mixed-gender public goods games, as opposed to when women only played the game. In hindsight, the authors reflect that the mixed-gender public goods game likely mimicked a community development council (CDC) -type situation (Fearon, Humphreys, and Weinstein 2011). One author suspected that if they were to use the RAPID test from DRC in Liberia, they would not have found the same positive results.

It is also sometimes difficult to accurately interpret games and other behavioural measures. For example, in one structured community activity in Sierra Leone, communities were given a tarp; the tarp ending up in the home of a leader would be interpreted as elite capture. But, an alternate interpretation arising in at least one interview is that the community didn't know what to do with such a gift, so stored it in a village elder's back room. In Aceh, as a behavioural measure, respondents were given about $1US and asked to decide in private how much they wanted to contribute to development activities by district and how much they wanted to keep for themselves. The study authors write that "the share of funds sent to local government serves as a measure for trust in the ability or willingness of local government to use funds well" (Barron et al. 2009 58). It could also indicate a number of alternatives: altruism, for instance. Moreover, the amount of money offered may not be enough to incite the kind of decision-making that greater sums may.

Are the evaluations focusing on the right level of analysis? Since leaders are exposed to much more intensive facilitation in most of the programmes under study here, one can expect different changes at the level of leadership and that of the general population. As the DRC report suggests, "it is possible that the primary effects of *Tuungane* are on leaders in communities, for example those that took part in trainings directly. If this is the case, the research is not well calibrated to capture these effects" (Humphreys, Sanchez de la Sierra, and van der Windt 2012, 75). At present, these programmes are targeted at changing general population attitudes and behaviours and the evaluations are measuring this accordingly.

Related, **the focus on average treatment effect may mask important information.** The idea of an average treatment effect is to estimate the average effects across a lot of heterogeneity. This was the goal of the evaluations included here. Nonetheless, a number of the studies note heterogeneous treatment effects. For instance, in Aceh, there were often different results for victims alone versus the whole sample. In Liberia, the evaluation revealed different results in the public goods game for mixed-gender and women only games. As the authors of the DRC study write, responding to their findings entailing mostly no or weak effects, "it is… possible that there

are positive effects for some and negative effects for others, with small or no effects on average. For example there may be differences related to prior levels of poverty or exposure to conflict, or social outcomes could depend on the success of the economic projects introduced. While our analysis in this report does not examine such heterogeneous effects, further research in these directions may help contribute to programme design decisions and help in assessment of the extent to which the limited effects of CDR identified here are general" (Humphreys, Sanchez de la Sierra, and van der Windt 2012, 74-5). Indeed, according to interviews, programme implementers in various countries heard from staff on the ground that you could see and feel differences in project communities, indicating that the project was having some positive effect. These staff wondered how it was that this was not always captured in the evaluations.

Is the timing of measurement of outcomes correct? It is possible that economic effects will not appear within the timeline of these evaluations. For example, if classrooms were constructed, and schooling attendance or quality increases, the economic return of increased access to education would not be witnessed for many years. It may well, for instance, make sense to invest in a later round of evaluation for Tuungane (Humphreys 2012). It is unlikely, however, that social cohesion or governance effects will increase after the evaluation with no additional intervention. On these effects, one would expect them to be strongest during and immediately after the project and to decrease with time.

How did the evaluation strategies impact the projects? Evaluation designs affected the projects in several ways. A first example relates to the selection of programme areas. An RCT requires that communities are selected into the project randomly. To make the lottery possible, this involves choosing twice as many eligible communities as those that would ultimately receive the project. This could mean that treatment communities are more spread out than they would otherwise be, resulting in cost increases for the programme. Another example relates to timing. Some interviewees suggested that set timelines for evaluations are pushing interventions to go faster in order to "get things done" in time for the evaluation. The effect can work the other way as well. In DRC, the evaluation focused on Tuungane phase 1 and resulted in delaying Tuungane phase 2 until the evaluation was complete. Finally, interviewees sometimes raised concerns that the evaluation pulled resources away from implementation. Several noted the importance of finding the right balance between implementation and evaluation.

Did the evaluations themselves have impacts on outcomes? In some cases, the evaluation may have impacted outcomes. When a research team spends a lot of time in the communities, or the community knows it is being monitored, the project and research become more entangled. In Afghanistan, for instance, the block grant disbursements in the districts they were studying were quite timely, in contrast to other areas, not being monitored, where grants were reportedly disbursed more slowly.

In some cases, randomization itself also impacted outcomes. In Liberia and Sierra Leone communities liked the public lotteries as transparent and fair (Fearon, Humphreys, and Weinstein

2008 34; Sulley et al. 2010, 19-20). In the case of Sierra Leone, though, "some of the selected villages have been privileged by the system… Some communities are neither 'villages' by GoBifo standards nor are they deprived communities" (Sulley et al. 2010, 19). But, "had the project tried to apply any criteria to the list before the lottery took place there would have been a risk of politicising the process" (Sulley et al. 2010, 20). Some of the staff that worked on the project reported that randomization reduced the individual patronage pressures they usually felt from family and friends.

In addition to challenges implementing CDD/R projects, there are **difficulties in implementing evaluations**. These include accessing difficult to reach, often unsafe, locations. As the research team from DRC reports, "the harsh conditions produced great costs to enumerators with high incidence of sickness including malaria and cholera. Although safety regulations were in place in all areas, one of the teams was involved in a tragic accident in which a child died" (Humphreys, Sanchez de la Sierra and van der Windt 2012, 34-5). Over the course of the evaluation in Afghanistan, some of the communities that were relatively safe at the beginning of the evaluation became very dangerous to visit by the endline. There are also practical difficulties, such as supervising data collection. To address potential concerns of enumeration fraud and their inability to directly supervise many of the survey activities, the Afghanistan team insisted that enumerators carry GPS tracking devices that reported their locations in real-time, as well as by more conventional methods. Although this created some friction, it allowed the team important checks on quality. A related problem caused by the high costs of local skilled labor as well security constraints was the difficulty of undertaking thorough qualitative work to accompany quantitative studies in some contexts.

Did the evaluations adequately explore unanticipated or adverse outcomes? The studies under consideration here raised several tensions worthy of further exploration. For instance, the Afghanistan and Aceh projects created some tension within communities due to the mandated female participation requirement which goes against cultural norms (Morel, Watanabe, and Wrobel 2009 24). In Afghanistan, this sometimes resulted in women "being denied meaningful participation in the programme, despite a strong interest, and projects favored by or benefiting women getting de-prioritized by powerful male elites" (Beath, Christia and Enikolopov 2012b, 8, citing Boesen 2004; Brick 2008). In Aceh, the programme's explicit focus on identifying and benefiting conflict victims was a controversial and sensitive issue. Some ex-combatants were frustrated to be excluded. When communities decided to include them to avoid social tension, conflict victims sometimes felt bitter (Morel, Watanabe, and Wrobel 2009 17-19, 27; Barron et al. 2009 29-33). As Barron et al, write "CDD programmes, including those with peacebuilding aims, can lead to tensions between groups through their promotion of competition over finite resources. In the long run, they may lead to a stronger basis for peace, through empowering groups and building local institutions. But in the short run, they can lead to social divisions.

Weighing these (potential) short and long run impacts is important in post-conflict environments" (2009, 68).

The responsibility to "do no harm" (Anderson 1999) prompts difficult questions for programming and evaluation. A number of reviewers were concerned that by phrasing hypotheses in the positive and looking predominantly for positive effects, the studies may have marginalised other unintended or adverse effects. Others raised the potential concern that experimental research, when inadequately paired with solid process evaluation or in-depth qualitative research, may miss or inadequately explore additional adverse effects.

While a thorough answer lies beyond the scope of this review, one can think about **how additional and sometimes different evidence may have been generated by complementary methodological approaches**. For example, the evaluations cited here relied primarily, but not exclusively, on quantitative data. As Betsy Levy-Paluck explains, it is "not that qualitative measurement would have made the experimental results "richer" or more detailed, although that is certainly the case. Using qualitative research methods in [a] field experiment could have provided a different understanding of the causal effect, identified possible causal mechanisms of change, and framed new interpretive understandings of [such issues as] authority, democracy, and gender within an experimentally assessed instance of social change"(2008, see also Barron, Diprose, and Woolcock 2011). **Different questions**, about programme design, implementation and process, for instance, **could also have been asked and some best answered with other methodological approaches**.

Is it context?

The moderating factors in the theory of change above –conflict, social cohesion, governance, and economic well-being – testify to the importance of context. Each of these factors is likely to affect the success of CDD/R. It was a driving principle of this review to separate out CDD/R in conflict-affected contexts from other CDD programmes; there are conflict-specific challenges to programme implementation and evaluation and it is thought that the conflict-affected nature of the contexts under study will affect results. Yet, we still don't know enough about how the key moderating factors matter.

Indeed, a key finding of this review is that **baseline measures of governance and social cohesion** – presumed to be low in conflict-affected contexts and part of the justification for CDD/R programming – **were often quite high**. In Liberia, for example, 90% of respondents belong to at least one social organization and 50% belong to at least three (Fearon, Humphreys, and Weinstein 2011 11). In the public goods game, nearly two thirds, even in control communities, made the maximum contribution (ibid, 26). These figures run counter the assumptions of significant intra-community tensions that guided programme design in Liberia (ibid). In DRC, the likelihood of using elections in the RAPID non-conditional cash transfer exercise was equally high in treatment and comparison communities – at approximately 50%

(Humphreys, Sanchez de la Sierra, and van der Windt 2012, 7). Even without a gender parity requirement, women comprised about 30% of committee members (ibid 74). In Aceh, social acceptance of ex-combatants and IDPs was already high, at 77% and 68%, respectively, in the control communities (Barron et al. 2009 47-8). These findings warrant rethinking of contextual presumptions and considering the possibility of threshold levels or tipping points for outcome measures.

Furthermore, just as project outputs need to be better matched to expected outcomes, discussed above, **CDD/R programming and evaluations need to do more to match goals, theories and outcomes to context**. For example, many interviewees hypothesized that for CDD/R to work, it has to be supported by government. Mansuri and Rao deem the best approaches to participatory development those that work with a "sandwich", based on effective top-down and bottom-up support (Mansuri and Rao 2012, 287). Based on this logic, we would expect to see stronger results from the Afghanistan, Sierra Leone, and Aceh programmes, which were better linked with government. While the results are not consistently different than DRC and Liberia, with virtually no real links to government, indirect future results may still transpire from these former CDD/R projects in the event that they become part of more lasting governmental change rather than only a "project". As another example, effects on gender perceptions or democratic attitudes will likely be more difficult to achieve in contexts in which patriarchal authoritative relations have long dominated social and political life. Goals, theories, programmes, and measures need to be adjusted to accordingly.

Conclusion and the road ahead for CDD/R

Based on a review of rigorous impact evaluations from programmes in Afghanistan, DRC, Aceh, Liberia and Sierra Leone, and interviews with practitioners, policymakers and academics, the record of CDD/R in conflict-affected contexts is mixed and, on the whole, discouraging. The five evaluations under study were designed to focus principally on the average treatment effects of aggregated packages of CDD/R. From this perspective, the ability of CDD/R programmes to reach the ambitious goals of improving socio-economic welfare, governance and social cohesion was limited. Likely explanations include a panacea-type approach to goals paired with an underspecified theory of change to guide programming; a variety of programme design issues including the relatively short design of CDD/R projects, the small size of block grants, the limited reach of the projects, the menu restrictions on CDD/R programming, the limitations of social infrastructure, the quality and intensity of social facilitation, the manner in which communities are conceptualised and thus often not meaningful to participants, and how community institutions build on existing institutions and relate to the state; and inadequate understanding of specific contextual factors and how they affect outcomes. These findings inform the road ahead for CDD/R.

Moving forward, CDD/R can only improve if policy-makers, practitioners, researchers, donors, and community members are encouraged to be open about genuine opinions and critiques and act on them.

- Interviews revealed that there are private and public transcripts about CDD/R: opinions that people are willing to share about CDD/R in private and those that they share in public (Scott 1990; Cooke and Kothari 2001). This hinders progress.

- CDD/R is chosen for a confluence of different reasons; being clear and honest about these reasons is a first step towards improving programming and evaluation. Is it the belief in intrinsic value in participation? That there is instrumental value in participation? A belief that CDD/R is better than alternatives for producing results? Or, the relative speediness of this method? Is it the ability to work in environments with weak institutions, undergoing transition? Is it that it seems to be one intervention capable of multiple outcomes? Is it funders incentives or an ideological commitment? Is it presumptions about conflict-affected contexts that make CDD/R appealing?

- Open dialogue should acknowledge tensions among those involved in CDD/R and aim to understand goals, priorities and trade-offs and build relationships. For example, within the donor community, programme implementation organisations, and academia, there are evaluation specialists, subject matter or programming specialists, and specific geographic area specialists. Within each of these areas, people support different methodological approaches, have different questions about programmes, and think differently about CDD/R. There is usually inadequate dialogue across these various stakeholders. In relation to evaluations, there are tensions between people who want an evaluation and people who feel that they are being forced into an evaluation. Some believe that evaluations are best conducted by those closest to the project, whereas others contend that only independent evaluators can offer clear and unbiased answers to questions; still others emphasize cooperation between insiders and outsiders. There are also tensions between researchers who suggest specific evaluation methods and implementers who do not understand, believe in, or support the methods and the implications that come with them. The producers and users of research may also have different priorities.

- As Mansuri and Rao write, "there needs to be a tolerance for honest feedback to facilitate learning, instead of a tendency to rush to judgment coupled with a pervasive fear of failure. The complexity of participatory development requires a high tolerance for failure and clear incentives for project managers to report evidence of it. Failure is sometimes the best way to learn about what works. Only in an environment in which failure is tolerated can innovation take place and evidence-based policy decisions be made" (2012,

14). This tolerance needs to come from donors, as well as policy-makers, practitioners, and researchers.

The way forward must also be guided by humility and more realistic goals. Alongside high needs of conflict-affected contexts and implementing agency specialisations, goals are often influenced by donor funding criteria. This frequently results in overpromising, setting goals that match buzzwords. Paired with rigorous impact evaluation, focused on packaged interventions and final outcomes, this is a highly risky undertaking. When outcomes are unrealistic and inadequately matched to the intervention, impact evaluations are not likely to find positive results. A risk that funders may consequently decide not to fund these programmes ensues, yet may be unfounded. Programme teams should more clearly define project goals and specify a contextually-specific theory of change with clear and realistic pathways to achieve them. They should work together to set mutually agreed upon goals that will survive individuals.

More questions can and should be asked of evaluations. Both literature and interviews identified consistent priority areas for a future research agenda on CDD/R.

- At present, the counterfactual to a CDD programme is having no project at all. This set of evaluations does not compare CDD to other types of projects, including top-down or centralized provision of services. To really answer questions for the aid community, and recognizing that there is no perfect way to deliver aid, we need to know which type of aid delivery works best. As Wong writes, "governments and the World Bank have very few tools to reach large numbers of the poor directly, particularly in the context of weak or fragile states, in post-conflict and post-disaster environments, or in areas with poor track records of service delivery within the bureaucracy" (2012, 4).

- CDD/R is currently implemented and evaluated as a bundled treatment – it has social (ie. facilitation and institution building) and economic (i.e. block grant) components. None of the evaluations to date can parse social from economic effects of the intervention. As the authors of the Sierra Leone evaluation note,"…for every dollar spent directly on community projects, roughly one dollar was spent ensuring the money was used well through facilitation, administration and oversight" (Casey, Glennerster, and Miguel 2011b, 11). They continue that "the key question this evaluation cannot address is whether the programme would be just as effective if the budget balance was shifted toward less facilitation and more grants to communities" (ibid, 45). In trying to decide how to enact change, and to best spend aid dollars, this is an important question.

- The evaluations included here all principally study the gross effects of a CDD/R programme; some describe them as "black box" experiments of CDD. The next wave of research could address any of the questions arising in the "is it the programming?" section above. For instance, the next wave of research could vary, within treatment communities, such programme features as the length and intensity of the interventions,

the scale of the investments, menu restrictions for projects, types of facilitation, different definitions of community and community institutions' ties with government. Contextual questions could also be answered through this type of variation. For instance, research could endeavour to parse the effects of CDD/R in areas with higher and lower baseline social cohesion to examine the importance of pre-existing social cohesion.

- More how and why questions need to be asked of the CDD/R process and the outcomes it generates. This line of inquiry could use qualitative methods, drawing on anthropological and broader ethnographic traditions, to gather detailed, in-depth, data about CDD/R interventions, the places in which they are implemented, and the processes they invoke. Indeed, such other types of research should meaningfully complement quantitative, experimental evaluations.

- Related, to move forward with evidence-based decision making, **implementers also need to take monitoring "far more seriously"** (Mansuri and Rao 2012). Most CDD/R projects today do not have effective monitoring or systems of learning. When implementation does not go in accordance with plan, programming and expectations are rarely adjusted. Interviews also revealed there is high turnover resulting in significant loss of institutional memory. In-depth process tracing could answer questions about trajectories and timelines. Ongoing process monitoring may also allow organisations to gather more evidence on the unexpected to help them "do no harm" – a heightened concern in conflict-affected contexts – and learn and adapt moving forward.

- As more studies emerge, and as datasets may become public, **statistical meta-analysis should also be possible**. It will need to be decided whether CDD/R in conflict-affected contexts is the most appropriate set of cases to compare, or if there are other groupings (i.e. CDD/R in Africa vs. Asia), wider (i.e. participatory development more broadly) or narrower (i.e. CDD/R in ongoing unstable conflict situations), that could give more analytical leverage.

This review presented a rare opportunity to stop and think about CDD/R programming and rigorous evaluations in conflict-affected contexts. The road ahead must build on the important work undertaken so far and tackle the many questions raised here, not simply replicate what has been done in the past.

Works Cited

3IE. 2012. *Checklist for making judgements about how much confidence to place in a systematic review of effects (adapted version of SURE checklist).* New Delhi: International Initiative for Impact Evaluation.

———. n.d. 3ie *Impact Evaluation Practice: A Guide for Grantees*. New Delhi: International Initiative for Impact Evaluation.

Anderson, Mary B. 1999. *Do no harm: how aid can support peace – or war*. Boulder: Lynne Rienner.

Arcand, Jean-Louis, Adama Bah, and Julien Labonne. 2010. *Conflict, Ideology and Foreign Aid Etudes et Documents,* E. 2010.21 (pp. 1-33).

Autesserre, Séverine. 2010. *The Trouble with the Congo: Local Violence and the Failure of International Peacekeeping*. New York: Cambridge University Press.

Banerjee, Abhijit V., and Esther Duflo. 2011. *Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty*. New York: PublicAffairs.

Barron, Patrick, Macartan Humphreys, Laura Paler, and Harvey M. Weinstein. 2009. "Community-Based Reintegration in Aceh: Assessing the Impacts of BRA-KDP,*"* in *Indonesian Social Development Paper No. 12*, eds. P. Barron. Jakarta: The World Bank.

Barron, Patrick. 2010. *CDD in Post-Conflict and Conflict Affected Areas: Experiences from East Asia In World Development Report 2011 - Background Report.* Washington DC: The

Barron, Patrick, Rachel Diprose, and Michael Woolcock. 2011. *Contesting Development: Participatory Projects and Local Conflict Dynamics in Indonesia.* New Haven: Yale University Press.

Beath, Andrew, Fotini Christia, and Ruben Enikolopov. 2011. Elite Capture of Local Institutions: Evidence from a Field Experiment in Afghanistan, (pp.1-46).

——— 2012a. Direct Democracy and Resource Allocation: Experimental Evidence from Afghanistan In *Policy Research Working Paper 6133*, edited by T. W. Bank. Washington DC: The World Bank.

———2012b. Empowering Women: Evidence from a Field Experiment in Afghanistan. *MIT Political Science (Working Paper No 2012-13).*

———2012c. Randomized Impact Evaluation of Afghanistan's National Solidarity Programme - Final report (Pre-Analysis Plan: Hypotheses, Methodology and Specifications).

———2012d. Winning Hearts and Minds through Development: Evidence from a Field Experiment in Afghanistan MIT Political Science (Working Paper):1-36.

Boesen I. 2004. "From Subjects to Citizens: Local Participation in the National Solidarity Programme," *AREU Working Paper Series*.

Bouvy, Alexis, and Maria Lange. 2012. *Ending the Deadlock: Towards a New Vision of Peace in Eastern DRC.* London: International Alert.

Brick, J. 2008. *Investigating the Sustainability of Community Development Councils in Afghanistan.* Kabul: Afghanistan Research and Evaluation Unit.

Casey, Katherine, Rachel Glennerster, and Edward Miguel. 2011a. *Reshaping Institutions: Evidence on Aid Impacts Using a Pre-Analysis Plan* (pp. 1-43).

———2011b. "Reshaping Institutions: Evidence on Aid Impacts Using a Pre-Analysis Plan," in *NBER Working Paper Series*, Cambridge: National Bureau of Economic Research.

———2011c. The GoBifo Project Evaluation Report: Assessing the Impacts of Community Driven Development in Sierra Leone.

Cliffe, Sarah, Scott Guggenheim, and Markus Kostner. 2003. "Community-Driven Reconstruction as an Instrument in War-to-Peace Transition," in *CPR Working Papers*. Washington DC: The World Bank.

Cooke, Bill, and Uma Kothari. 2001. "The Case for Participation as Tyranny," in *Participation: The New Tyranny?,* eds. B. Cooke and U. Kothari. London: Zed Books.

Dennys, Christian, and Idrees Zaman. 2009. *Trends in local Afghan conflicts: Synthesis paper.* Kabul: Cooperation for Peace and Unity.

Dongier, Philippe, Julie Van Domelen, Elinor Ostrom, Andrea Rizvi, Wendy Wakeman, Anthony Bebbington, Sabina Alkire, Talib Esmail, and Margaret Polski. 2001. "Community-Driven Development," in *Poverty Reduction Strategy Paper Sourcebook.* Washington DC: The World Bank.

Earth Institute, at Columbia University. 2008. *Millennium Villages: A New Approach to Fighting Poverty 2008* [cited October 8 2008]. Available from http://www.millenniumvillages.org.

Edillon, Rosemarie , Sharon Faye Piza, and Carlos Abad Santos. 2011. *Final Survey for the Kalahi-CIDSS Impact Evaluation,* edited by A. P. P. Center. Washington DC: The World Bank.

Fearon, James D., Macartan Humphreys, and Jeremy Weinstein. 2008. *Community-Driven Reconstruction in Lofa County: Impact Assessment.* New York: Stanford & Columbia Universities.

——— 2011. *Democratic Institutions and Collective Action Capacity: Results from a Field Experiement in Post-Conflict Liberia.* APSA 2011 Annual Meeting Paper, (pp.1-48).

Foreign Policy, and Fund for Peace. 2012. *Failed States: The 2012 Index.* Foreign Policy.

Haider, Huma. 2009. *Community-Based Approaches to Peacebuilding in Conflict-Affected and Fragile Contexts (Issue Paper).* Governance and Social Development Resource Centre.

Humphreys, Macartan. 2012a. CDR Evaluation Evidence, Implications, Method (Presentation). London: DFID.

——— 2012b. Evaluation: RCTs and More (presentation). Nairobi.

Humphreys, Macartan, Raul Sanchez de la Sierra, and Peter van der Windt. 2011. *Tuungane 1: Outcomes and Data Sources.*

——— 2012. *Social and Economic Impacts of Tuungane: Final Report on the Effects of a Community Driven Reconstruction Programme in Eastern Democratic Republic of Congo.*

Hussein, Adam M. 2004. "Somalia: International Versus Local Attempts at Peacebuilding," in *Durable Peace: Challenges for Peacebuilding in Africa.* Toronto: University of Toronto Press.

International Development Association. 2009. *Community-Driven Development: Delivering the Results People Need.* Washington DC: The World Bank.

International Rescue Committee. Undated. *Community-Driven Reconstruction Factsheet.* New York & London: Governance and Rights Technical Unit.

International Rescue Committee. 2006a. *Communities Sowing the Future of Congo.* London: The International Rescue Committee.

———. 2006b. *Community-Driven Reintegration and Reconstruction in Rural Liberia.* Monrovia: The International Rescue Committee.

———. 2006c. CDRR LOG FRAME

———. 2008. *Community Driven Reconstruction and Reintegration in Liberia: An Internal Process Evaluation.* The International Rescue Committee.

Kalyvas, Stathis N. 2006. *The Logic of Violence in Civil War.* Cambridge and New York: Cambridge University Press.

Karlan, Dean, and Jacob Appel. 2011. *More Than Good Intentions: How a New Economics Is Helping to Solve Global Poverty.* New York: Dutton.

King, Elisabeth, Cyrus Samii, and Birte Snilstveit. 2010. "Interventions to Promote Social Cohesion in Sub-Saharan Africa". *Journal of Development Effectiveness*, 2 (3): 336-72.

Kumar, Nalini. 2005. *The Effectiveness of World Bank Support for Community-based and -driven Development: An OED Evaluation.* Herndon, VA: World Bank.

Labonne, Julien, and Robert S. Chase. 2008. *Do Community-Driven Development Projects Enhance Social Capital? Evidence from the Philippines.* Washington: World Bank - Social Development Department.

Leonard, David K. 2013. *Piecing it Together: Post-Conflict Security in Africa of Networked, Multilevel Governance*. Brighton: Institute of Development Studies.

Levy-Paluck, Elizabeth. 2008. "The Promising Integration of Qualitative Methods and Field Experiments." *Newsletter of the American Political Science Association Organized Section for Qualitative and Multi-Method Research*, 6 (2):23-29.

Mansuri, Ghazala, and Vijayendra Rao. 2004. "Community Based (and Driven) Development: A Critical Review." *World Bank Policy Research Working Paper No. 3209*

———2012. *Localizing Development: Does Participation Really Work?* Washington DC: World Bank.

Maynard, Kimberly. 2007. *The Role of Culture, Islam and Tradition in Community Driven Reconstruction: A Study on the IRC's Approach to Afghanistan's National Solidarity Programme*. The International Rescue Committee.

McBride, Lizanne, and Naina Patel. 2007. *Lessons Learned on Community Driven Reconstruction*. The International Rescue Committee.

McBride, Lizanne, and Alyoscia D'Onofrio. 2008. "Community-Driven Reconstruction: A New Strategy for Recovery," *Humanitarian Exchange Magazine* (39).

Morel, Adrian, Makiko Watanabe, and Robert Wrobel. 2009. Delivering Assistance to Conflict-Affected Communities: The BRA-KDP in Aceh," in *Indonesian Social Development Paper No. 13*. Jakarta: The World Bank.

Narayan, Deepa, R. Patel, K. Schafft, A. Rademacher, and S. Koch-Schulte. 2000. *Voices of the poor: can anyone hear us?* New York: Oxford University Press, for The World Bank.

Olken, Benjamin A., Junko Onishi, and Susan Wong. 2011. *Indonesia's PNPM Generasi Programme: Final Impact Evaluation Report*.

Peace Direct, and Quaker United Nations Office. 2010. *Ripples into Waves: Locally Led Peacebuilding on a National Scale*. New York: Peace Direct.

Scott, James. 1990. *Domination and the Arts of Resistance: Hidden Transcripts*. New Haven: Yale University Press.

Sen, Amartya. 1999. *Development as Freedom*. New York: Anchor Books.

Slaymaker, Tom, Karin Christiansen, and Isabel Hemming. 2005. *Community-based approaches and service delivery: Issues and Options in Difficult Environment and Partnerships*. London: Overseas Development Institute.

Social Development Department. 2006. *Community-Driven Development in the Context of Conflict-Affected Countries: Challenges and Opportunities*. Washington DC: The World Bank.

Strand, Strand, Hege Toje, Alf Morten Jerve, and Ingrid Samset. 2003. *Community Driven Development in Contexts of Conflict*. Washington DC: The World Bank.

Straus, Scott. 2006. *The Order of Genocide: Race, Power, and War in Rwanda*. Ithaca: Cornell University Press.

Sulley, Sesay , Victor Amara, John Abraham Lebbi, Minkahil Bangura, Mohammed Mankay Sesay, Joseph Sivalie, John Alusine Kaloko, Augustine Moriba, Santigie Kany, Theophilus Dimoh, Abdul Karim Tarawalie, and Joseph Johnson. 2010. *Strengthening Social Capital: The GoBifo Approach in Rural Sierra Leone*, eds. G. Baltissen and M. Manyeh. KIT Publishers.

UNDP. 2011.*Sustainability and Equity: A Better Future for All.* New York: United Nations Development Programme.

USAID. 2007. *Community-Based Development in Conflict-Affected Areas: An Introductory Guide for Programming.* Washington DC: US Agency for International Development.

Vajja, Anju, and Howard White. 2008. "Can the World Bank build social capital? The experience of social funds in Malawi and Zambia," *Journal of Development Studies*, 44 (8), 1145–1168.

Voss, John. 2008. *Impact Evaluation of the Second Phase of the Kecamatan Development Program in Indonesia.* Jakarta: The World Bank.

White, Howard. 2013. *An introduction to the use of randomized control trials to evaluate development interventions.* New Delhi: International Initiative for Impact Evaluation.

Wong, Susan 2012. *What Have Been the Impacts of World Bank Community-Driven Development Programs? CDD Impact Evaluation Review and Operational & Research Implications.* Washington DC: The World Bank.

World Bank. 2006. *Community-Driven Development in the Context of Conflict-Affected Countries: Challenges and Opportunities.* Washington DC: The World Bank.