

Humphreys, Macartan; de la Sierra, Raúl Sánchez; der Windt, Peter Van

Article — Accepted Manuscript (Postprint)

Exporting democratic practices: Evidence from a village governance intervention in Eastern Congo

Journal of Development Economics

Provided in Cooperation with:

WZB Berlin Social Science Center

Suggested Citation: Humphreys, Macartan; de la Sierra, Raúl Sánchez; der Windt, Peter Van (2019) : Exporting democratic practices: Evidence from a village governance intervention in Eastern Congo, Journal of Development Economics, ISSN 0304-3878, Elsevier, Amsterdam, Vol. 140, pp. 279-301, <https://doi.org/10.1016/j.jdeveco.2019.03.011>

This Version is available at:

<https://hdl.handle.net/10419/214632>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Exporting Democratic Practices: Evidence from a Village Governance Intervention in Eastern Congo

Macartan Humphreys*

Raul Sanchez de la Sierra†

Peter Van der Windt‡

June 15, 2019

Abstract

We study a randomized Community Driven Reconstruction (CDR) intervention that provided two years of exposure to democratic practices in 1,250 villages in eastern Congo. To assess impacts, we examine behavior in a later village-level unconditional cash transfer project that distributed \$1,000 to 457 treatment and control villages. The exercise provides opportunities to assess whether public funds get captured, what governance practices are employed by villagers and village elites and whether the intervention altered these behaviors. We find no evidence for such effects. The results cast doubt on current attempts to export democratic practices to local communities.

JEL Codes: D72; P48; D02; O17

Keywords: Political Processes; Political Economy; Institutions; Culture; Demonstration Effects

Forthcoming in *Journal of Development Economics*.

*Columbia University and WZB Berlin Social Science Center. mh2245@columbia.edu.

†Corresponding author. UC Berkeley. rsanchezdelasierra@berkeley.edu.

‡New York University - Abu Dhabi. petervanderwindt@nyu.edu. This research was funded by the International Initiative for Impact Evaluation (3IE) and the Department For International Development, UK. We thank the International Rescue Committee, CARE International, and Chimanku Bantuzeko for their partnership in that research. We thank Lily Medina and Clara Bicalho for excellent support with analysis. Humphreys thanks the Trudeau Foundation for support while this work was undertaken. This research was approved by Columbia University IRB and was implemented with consent of participating subjects. The study did not make use of deception at any stage. Van der Windt thanks Wageningen University. Hypotheses, econometric specifications and covariates were specified prior to the development of social science registries. See Humphreys et al. (2013); deviations from this plan are discussed in appendix L.

1 Introduction

Research in political economy and long-run development suggests that institutions are a key driver of economic development. Since economic and political institutions constrain incentives, they can produce variation in the extent to which allocations benefit populations or elites, and, ultimately, economic growth (Acemoglu et al., 2001). Although political institutions are likely to change slowly, international aid organizations and Western governments have taken a cue that they are important and have sought to change them in developing areas where they deem them weak. One common approach is to provide short-term exposure to the practice of institutions that are inclusive, or more democratic, in the hope that these will get picked up and lead to greater accountability of local elites.

A prominent version of this approach in post-conflict settings uses the “Community Driven Reconstruction” (CDR) model. CDR programs, a variant of Community Driven Development (CDD) programs, are part of the “participatory development” model for development — a popular model that accounted for \$85bn in World Bank spending in one decade alone (Mansuri and Rao, 2013). While many participatory development programs seek to use participation to enhance program effectiveness or program legitimacy, many CDR programs stand out for their transformative ambitions, seeking not just to involve community institutions, but to refashion them. However, the belief that this approach can improve institutions hinges on two questionable assumptions. First, that governance practices of conflict-affected areas need to be changed. And, second, that exposure to foreign governance practices — believed to be better for the population — will be adopted. While some studies suggest that CDR programs have an effect on subsequent public goods games for subgroups (Fearon et al., 2009), there is almost no evidence that CDR programs have any subsequent effect on governance practices (Casey et al., 2013; Beath et al., 2013; King and Samii, 2014).

In this paper, we study an unusually large randomized CDR program and ask whether it caused subsequent village allocations of public funds to be less captured by village elites and whether it induced more democratic governance practices. The program targeted 1,250 villages in the Democratic Republic of the Congo (henceforth, the Congo) to undergo democratic training and practice in the management of development funds over two years. It was implemented by two US non-governmental organizations (NGOs) operating in the Congo and had a total cost of \$46m and a target beneficiary population of 1,780,000 people. During the program, villages were trained in the implementation of elections, in accounting, and in accountability practices and were exposed to advocacy for democratic processes (through awareness raising campaigns where “good governance” practices were introduced). Participating populations then put these ideas into practice by selecting a development project for

the village and electing a management committee that managed project funds. Populations were tasked with holding the committee accountable through frequent town hall meetings. This intervention thus created exposure to the democratic process at the village level, which donors believed to be sufficiently prolonged to induce a change in local governance practices.

We explore the impact of the CDR program along two dimensions. We first focus on the allocation of public funds, particularly the capture of benefits by the elite. We then focus on three dimensions that measure the degree to which governance practices are democratic: inclusiveness of the process (participation), accountability of the elite (accountability), and transparency of the elite actions (transparency).

To measure the allocation of public funds and the degree to which governance practices are democratic, we observe community behavior in a real public funds allocation decision after about two years of CDR program exposure. To induce a comparable public funds problem in each village, we provide village-level unconditional cash transfers of \$1,000 to 457 villages, half of which had been previously treated by the CDR program, the other half not. We use this cash transfer project (dubbed “RAPID”) to observe how the community solves the public funds allocation problem, and the governance practices that it sets in motion. The RAPID cash transfer project was implemented in four steps and took two to three months in each community. Communities were free to use the funds as they chose and to decide how to manage the use of funds. The project was designed so that we could measure changes in the allocation of public funds and governance practices, using direct observations, surveys conducted in private at the start and end of the project to a random sample of households and to the chief, focus groups with villagers and the elite, and a comprehensive audit. We focus on transparency of information held by elites, participation of villagers, composition of the committee and kinship relations, funds misuse and corruption, predominance of villagers and chief’s preferences in influencing the public funds allocation in the RAPID cash transfer project.

Our first result is that, despite the scale and duration of the CDR intervention, we find no evidence of impact on the extent to which there is capture of benefits by the elite, as measured by the allocation of benefits from the RAPID cash transfer program.

Our second result is that the failure of the CDR program to affect elite capture is accompanied by a failure to affect democratic practices along the three dimensions that we measure. In particular, the program leaves unchanged the patterns of inclusiveness and participation in the governance processes in the village, the degree to which the population holds the elite accountable, and the level of transparency regarding the usage of public funds in the village.

Our third result is that these findings cannot be easily explained by low statistical power, poor CDR program implementation, spillovers to control areas, social desirability in control areas, poor measurement, elite backlash against democratization, or changes in expectations about future aid in control villages. Our findings are thus a strong indication that adoption of democratic institutions does not follow from two years of training and practice.

Eastern Congo is a well-suited environment to examine the adoption of democratic practice in local governance. The state has largely withdrawn from the rural areas of the east and enjoys low legitimacy. Local governance is often described as “captured” by traditional chiefs and vulnerable to corrupt practices by state officials. These features are not unique to the Congo. Multiple accounts suggest that in many Sub-Saharan states, colonial rule used pre-colonial institutions to create “decentralized despots” in ways that are detrimental to development (Acemoglu et al., 2014a; Mamdani, 1996). The types of democratic practices introduced by this program are believed to be largely new in the rural areas, where chiefs inherit power through lines of succession and where chief accountability hinges on other outcomes than is usually assumed by Western donors, practitioners, and academics (Hoffmann, 2014).

This study complements political economy and development literatures that study the role of institutions. The view that liberal democratic institutions, inspired by Western models, are more conducive to growth finds support in academic research (Acemoglu et al., 2001; Acemoglu and Robinson, 2012). Furthermore, our study has implications for development policy, at a moment where billions are being spent by Western donors and international organizations in transforming institutions with little evidence to support the methods or the objectives of such an approach (Mansuri and Rao, 2013).

In addition, our study makes three methodological contributions.

First, we develop a novel measure of governance practices that hinges on observation of behavior of typical households and elites in order to obtain a comprehensive characterization of the social process that underpins them. By introducing real economic trade-offs, this paper’s approach to measurement improves upon standard self-reported measurements, which are subject to reporting and desirability bias (Barron et al., 2009). It also improves upon laboratory games, where interpretation of effects can be rendered difficult by the absence of natural metrics or sensitivity to features of artificial environments (Fearon et al., 2009; Lowes et al., 2017; Haley and Fessler, 2005; Avdeenko and Gilligan, 2015). Although our strategy is similar to some of the activities studied by Casey et al. (2013) and Beath et al. (2013), our measurement strategy complements these studies in that the behaviors that we observe occur in a large-stakes environment, in a forum with minimal control by the researchers, and

over an extended period of time. It is thus more straightforward to interpret our results as “natural” governance processes and outcomes of the community. While the large number of outcomes introduces multiple comparisons problems, which the analysis addresses, it also allows us to characterize governance practices more comprehensively than previous studies.

Second, we address the concern that the absence of effects identified in this and other studies could arise because of positive spillovers to control areas. This is a particularly important concern, because the treatment is basically information and practice, which can easily spread through social networks, especially if beneficiaries think it is a positive adoption. Indeed the underlying theory behind these interventions assumes that exposure induces take-up. However, to date, no study of CDR or CDD has tackled the issue of spillovers. Using a model of indirect effects, our approach allows us to examine spatial spillovers in a way that makes use of exposure propensities implied by the randomization procedure and that hinges on relatively few assumptions (see Aronow and Samii (2013)).

Third, to our knowledge, our study provides the results of the largest CDR program experiment to date. The CDR program was implemented in 1,250 villages across 560 village clusters. Our measurement strategy employs data collected in 457 village clusters, significantly reducing the likelihood of false negatives. As a comparison, Fearon et al. (2009) examine 83 villages, Casey et al. (2013) examine 236 villages, and Beath et al. (2011) examine 217 village clusters (Casey, 2018).

Section 2 discusses the intervention we study in the context of the literature and related theory. Section 3 anchors the study in the Congolese context. Section 4 presents the details of the intervention. We present the empirical strategy and the results in Sections 5 and 6. Section 7 discusses whether the null effect is due to weaknesses of the experimental design. We conclude in Section 8.

2 Research Question and Existing Literature

We examine an ambitious implementation of the CDR model. The CDR program that we examine, supported by the UK Department For International Development (DFID) with a budget of \$46 million, was one of the largest programs of its form. This led the CDR program to be called “one of the world’s largest ever randomized trials” (see Hartford (2014) and Hartford (2012)). Donors were enthusiastic about the transformative effects of this CDR program on governance culture in the villages targeted. From data we gathered before the start of data collection, most decision-makers at the implementing agency thought it possible or very likely that the CDR program would have an impact on each dimension of

governance. Half of them thought it very likely that CDR program villages would become more transparent and inclusive.¹

Other studies also examine whether participatory development can alter governance practices (see especially reviews by King and Samii (2014) and Casey (2018)). We differ from these studies in the following ways. First, while Fearon et al. (2009) examine whether bringing people together increases the valuation of the public good, our study focuses instead on whether democratic governance practice can be adopted, leading to a change in political institutions. Second, while Beath et al. (2013) examine whether state aid increases state legitimacy by creating reciprocity, we focus instead on external aid as the vehicle of transmission of governance practices. This distinction is also policy relevant, since aid is often a promising lever of change when the state is ineffective.

Our study is closest to Casey et al. (2013), but complements it in a number of ways, besides providing variation in location and scale. While Casey et al. (2013) examine a program that “reconstitutes elected district-level governments,” we examine a CDR program that introduces exposure to parallel, and new, democratic institutions and aims to induce change in governance practice rather than reinforce the capacity of recently created institutions of governance. The intervention in Casey et al. (2013) aims to promote the effectiveness of institutions that are part of an administrative system that already exists — created just after the war — and that will continue to exist after the program, trying to foster participation and inclusion. The phase of the program we study also reflected the desire to focus on institution-building rather than infrastructural development. As we describe below, only \$16m out of the \$46m budget in the CDR program we study went towards the actual infrastructure projects, compared to an approximate one-to-one allocation in Casey et al. (2013), and the facilitation costs were front-loaded for the first stage of the program. Finally, there are important differences in measurement strategy. Casey et al. (2013) use innovative structured exercises to observe behavior, such as decisions over the choice of a battery or salt in the presence of enumerators or the use to which a gift of tarp was put. In contrast, our strategy sought to emulate a realistic and unstructured project management problem with minimal control over what choices should be made or what institutions should be used.

In the appendix (A), we also discuss how this study relates to more theoretical literature, illustrating the mechanisms through which exposure to democratic practices might alter political institutions.

¹Prior to launching our endline data collection, we conducted a survey with CDR program implementers and program directors (12 respondents) and seven researchers working in the region. Researchers, in contrast, were considerably more skeptical that traditional leaders would become more accountable.

3 Governance Practices in Eastern Congo

The donors that supported the CDR intervention aimed to change local governance practices in eastern Congo. We provide here a short description of pre-existing governance practices to clarify that the intervention we examine took place in a setting where institutions of governance already existed and were entrenched, rather than in one in which institutions were destroyed and needed to be reconstructed.

Public authority in rural areas of eastern Congo is mostly embedded in customary chiefs (Newbury, 1991), which are positions of informal power that predate the colonial state. In addition, in the North and South Kivu provinces particularly, non-state armed organizations collect regular taxes, provide protection, and run fiscal administrations (Hoffmann et al., 2016; Stearns et al., 2013; Stearns and Vogel, 2015; Raeymaekers, 2014; Sanchez de la Sierra, 2017).

Customary chiefs can be one of two types: a village chief or a chiefdom chief, the “Mwamis.” Mwamis control a larger area, called “chefferie,” which is composed of tens to hundreds of villages. Village chiefs derive their power, in principle, from a set of governance practices that have been in place for generations (Akyeampong et al., 2014). By coordinating expectations, these practices are embedded with certain forms of legitimacy (Hoffmann, 2014; Newbury, 1991). Village chiefs are often enthroned following kin based lines of succession. When a village chief is enthroned, the chiefdom chief, with the help of local witch doctors, invokes the tribal ancestors to confirm the legitimacy of the new chief. Once enthroned, a village chief usually governs for life. The colonial state re-inforced the power of village as well as chiefdom chiefs. Indeed, the Belgian administrators co-opted customary chiefs, and obtained taxes, labor, and other resources through them, in exchange for the support of the coercive apparatus of the colonial state (Hoffmann, 2014; Acemoglu et al., 2014b; Mamdani, 1996). After independence, village and chiefdom chiefs remained as a basis for public authority.

Village chiefs are the owners of land which, according to custom, is where ancestors are buried, and which they can allocate to households in exchange for a tax. Village chiefs most often administer justice and taxation in the village. They also organize the provision of public goods (clearing the road, building infrastructure, and mobilizing self-defense groups), drawing on an old tradition of forced labor, *Salongo*. The power of village chiefs can hinge on perceptions of their supernatural talent, their toughness, and their leadership skills (Newbury, 1991).

4 Intervention: Community Driven Reconstruction

Our study takes advantage of a large UK funded CDR program, called “Tuungane,” implemented by the International Rescue Committee and CARE International in 1,250 villages throughout eastern Congo. The program had as a central goal to “improve the understanding and practice of democratic governance.”²

The goal of the program reflects broader normative goals shared by many international organizations and Western aid donors to promote a more recognizable democratic culture of governance. As a reflection of this goal, the program structurally sought to minimize the influence of traditional leaders in collective decision making. Implementers emphasized the normative desirability of the practices. The protocols of the program specifies aims of “[improving] good governance: practices of the transparency, accountability, representation, participatory management, inclusion of all.”

Over a four year period, the program spent \$46 million of development aid, reaching approximately 1,250 villages and a beneficiary population of approximately 1,780,000 people. A large share of this funding was used for facilitation and indirect costs, with only \$16m, 35% of the total program costs, going directly towards infrastructure. These shares reflect the fact that the main focus of the intervention was institutional change, not the use of existing institutions to deploy development funds. Because this study focused more on learning about the social impacts (rather than the economic impacts) of the program, we focus on the first two of the overall four years, a period in which the institutional components were frontloaded and only \$3.7m was spent on small-scale implementation. Figure 2 in the appendix (B) illustrates the timing of implementation (and data collection) across areas.

The program followed well defined steps. First, populations were mobilized to townhall meetings, where the objectives, the implementation agencies, and the funding government were introduced. Second, the population was trained to participate in local democracy, which many had never experienced before (14% of the chiefs in our sample are elected through elections). Third, members of the village were encouraged to run for these elections.³ Fourth, (private) elections were organized to elect a committee, whose task would be to manage the aid fund.⁴ Fifth, committee members submitted a proposed spending plan for popular

²In 2007, in collaboration with the implementing partner, the research team developed hypotheses that took account of these goals. A broader set of hypotheses relating to behavioral outcomes were developed during implementation and prior to data collection.

³The only requirement to run was: “People may nominate themselves, but if they do so, they are required to have at least two other people support them.” Source: Tuungane protocols.

⁴For elections to be valid, at least 70% of the adult population had to vote. “At least 70 percent of the adult (over the age of 18) voting population must vote in order for elections to be valid. It is the responsibility of the Election Team and Tuungane to ensure adequate participation. This is the only way to legitimize

approval in a village election.⁵ Sixth, the committee was, in principle, held to account by the population. On average, four general assemblies were convened by the committees to justify the use of program funds to populations. Elections were held where committee members could be revoked for funds embezzlement. Seventh, committee members received intense training in leadership, principles of governance and accountability, as well as financial management and accounting practices. Finally, to ensure the engagement of the average villager, communities contributed to their chosen community spending plan with cash or in-kind support, which aimed to create an ownership effect and thus create an additional link of accountability.

Records and audit reports from the implementing partner suggest that the program was technically well implemented. Meetings were held, trainings and elections took place, and infrastructure was built. Indeed, according to interviews with program implementers, the program was not allowed to fail. This was because the NGO was responsible for fund use and because a positive outcome was considered a part of the treatment for encouraging take-up of the institutional practices.

Our own data broadly support this positive view. 62% of the population in the CDR program villages knew about the CDR program, and 39% of those knew who implemented it. Furthermore, 76% of committee members and 48% of village chiefs were able to guess the right size of the grant, although only 22% of the general population guessed the correct CDR amount. We also recorded attendance at CDR program meetings. We find that 36% of men and 23% of women reported attending some meetings associated with the CDR program. More than half of the chiefs interviewed reported attending some meetings, and 84% of Tuungane committee members reported attendance; the top 25% claiming to have attended more than four meetings. The median chief reported attending four meetings while the top 25% attended seven or more; the equivalent numbers for the committee members are 9 and 15 meetings. The overall knowledge and participation in the CDR program was therefore considerable, especially given the fact that the intervention was not a part of the formal government administrative structure and that it was operating in difficult-to-access and sparsely populated areas.

As points of contrast, approximately 50 to 100 people, in villages of about 2,700, attended community meetings in the KDP CDR program in Indonesia (Guggenheim et al., 2004). Bjorkman and Svensson (2009) report that “more than” 150 participants attended meetings in a related intervention, for communities of around 2,500 people. A World Bank report on

the elections and ensure that they are in fact a product of the community as a whole.” Source: Tuungane protocols.

⁵If rejected, it was subsequently altered until a majority approved it.

KDP Mandiri reported low meeting attendance rates, with women accounting for less than 1% of attendees (Basri et al., 2004, p5).

5 Empirical Strategy

This section describes the treatment assignment, and our strategy to measure outcomes and estimate treatment effects.

5.1 Assignment to Treatment

Communities were grouped into 600 village clusters, which were in turn grouped geographically into 83 randomization “lottery blocks.” Within each lottery block, clusters of villages were randomly drawn to achieve balance in treatment status.⁶ This approach helps improve balance between treatment and control by geographic features including remoteness, poverty, institutions, and social composition. In total, all 600 village clusters entered the lottery, 280 were selected for treatment and the remaining 320 were in control (see Figure 1).

Public lotteries were held to implement the randomization process for a number of reasons. First, they provide a form of informed consent on the part of communities. Second, there is transparency over the selection process and this reduces concerns that one community was being unfairly favored over another. In principle, though, public lotteries could lead to jealousy which could lead to bias in estimates if, for instance, control communities may have started performing better or worse as a result of not being chosen. Our investigations of perceptions of the lotteries suggest that this is not a concern.⁷

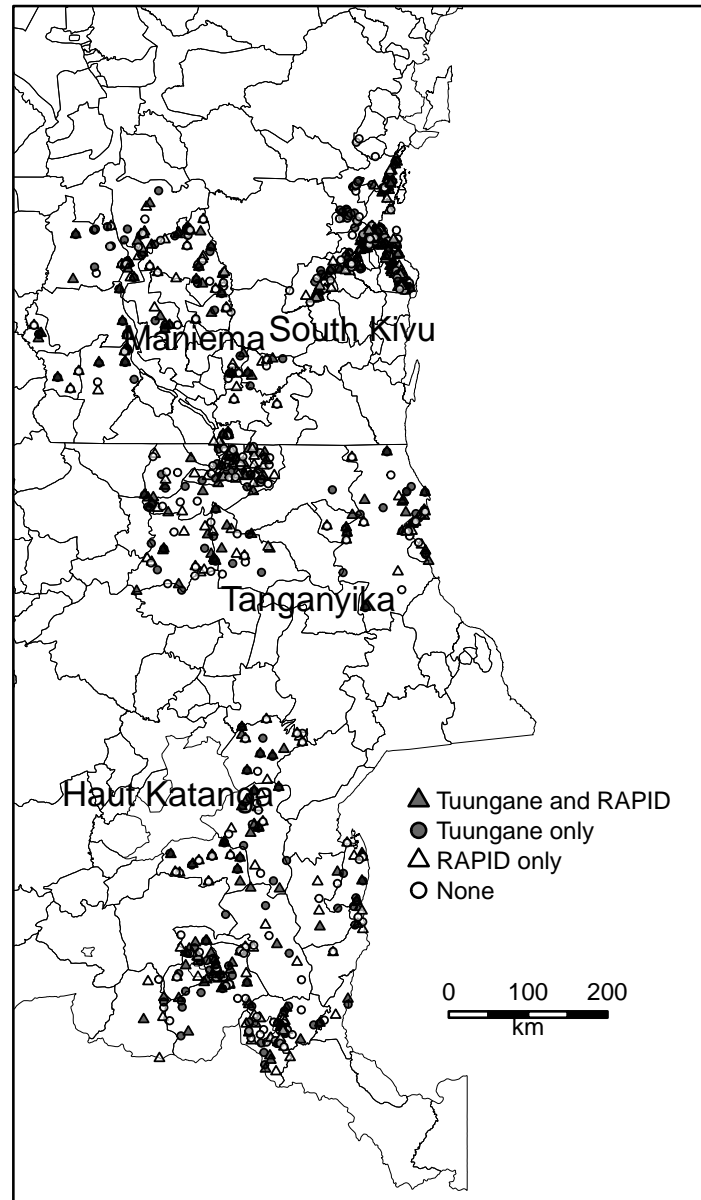
5.2 Measurement

In this section, we first describe the RAPID project, and then the measures that the project allows us to gather.

⁶Lottery blocks largely corresponded to chiefdoms (“Chefferies”) or sectors (“Secteurs”). For simplicity, we generally use the term chiefdom for both units.

⁷We asked a set of survey respondents (that had heard of Tuungane) in treatment and control areas how they thought communities were chosen. In treatment areas, 59% of those responding reported that the villages were chosen by chance. Divine intervention was the next most common answer. Few gave traditional explanations such as favoritism by government or NGOs. Patterns in control areas were largely similar although in these areas the vast majority of respondents either had not heard of Tuungane or had no explanation for why the program was not implemented in their community.

Figure 1: Distribution of Treatment and Control Clusters



Notes: This figure shows chiefdoms (polygons) and study villages in eastern Congo. Randomization was implemented at the level of blocks roughly corresponding to chiefdoms and ranging in size from 2 treatment units to 30 units. This figure maps 1,020 of the research villages for which we had GPS data. Source: study data.

5.2.1 The RAPID Project

After the first phase of the CDR program, we introduced an independent unconditional cash transfer project (“Recherche-Action sur les Projets d’Impact pour le Développement,”

henceforth RAPID) in treatment and control villages. Local university-managed teams delivered \$1,000 to consenting communities (consent was sought at village level separately for participating in the program and participating in the research, though in practice consent was never withheld). Communities were unconstrained in how they could spend the funds, except for a small set of prohibited uses.⁸ While the RAPID project moderately encouraged distributive spending, these were not required (the script read to the community during Step A is provided in the appendix (D)). Importantly, the unconditional cash transfer left communities free to decide who should manage the funds and how decisions should be made. We implemented the RAPID project in four steps (A-D) over the course of two to three months.

As part of the RAPID project we collected data from multiple sources. During Step A, a survey was conducted among five randomly selected villagers and the chief to learn about project preferences. Furthermore, as part of the village meeting that introduced the RAPID project we collect measures about the quality of participation. During Step B, we collect information about the village’s decisions regarding how to use funding and who is entrusted to manage it. As part of Step D, we undertake a carefully-designed audit and conduct focus groups to track the use of funds and learn about the RAPID implementation process. We also conduct surveys with ten randomly selected households, five of which were also surveyed during Step A. Table 1 gives an overview of the key features of the RAPID project, including the data collection activities.

RAPID was designed to limit the risk of desirability bias. The introduction of the RAPID program was not connected with the CDR program; there was no overlap in staff or branding. Moreover, the RAPID project funds were given unconditionally, and villages knew they were not rewarded for good behavior.

Due to budgetary constraints, we did not select all treatment and control villages of the CDR program to take part in the unconditional cash transfer project. Instead, we randomly sampled one village in all village clusters that received the treatment, and randomly sampled one village in the same number of village clusters in control areas. Since the CDR program was implemented across 280 village clusters, we thus targeted 560 villages for the unconditional cash transfer. We again used geographic blocks (corresponding to chiefdoms) to select units into RAPID (with additional blocking on participation in the CDR program). Thus, we ensure balance on geographic variables of RAPID communities.

The implementation of our measurement strategy was met by multiple practical challenges in the field, which led to attrition. In total, we were able to conduct RAPID’s Step A

⁸Some key constraints were that some uses were ruled out if this was likely to result in harm (such as the purchase of weapons) and the grant had to be spent out within a two month period – in order to be able to assess the use of funds in a timely manner.

in 457 villages, Step B and C in 454 villages and Step D in 413 villages. We collected Step A survey data from 2,214 households, and Step D survey data from 3,815 households, of which 1,863 are panel. The main reason for attrition were political tensions in the run up to the November 2012 presidential elections, which led to the expulsion of the Maniema teams. This loss covered entire lottery blocks, and is thus unrelated to the treatment status of units. We discuss attrition in detail in appendix K.3, which also presents a CONSORT-style flow chart with the details on the number of villages targeted and visited.⁹

In addition to these RAPID villages, we visited a second set of villages. That is, when we assigned villages to the RAPID project we also randomly selected an additional village in each village cluster to be a “survey-only” village. These additional 560 “survey-only” villages did not receive the RAPID program but were visited in parallel to RAPID’s Step D to collect survey data from five randomly selected households. In this study, only the behavioral measure of information seeking makes use of both RAPID and survey-only villages. Note that, because of attrition, we collected data in only 403 survey-only villages and from 1,929 households.

We describe now how the RAPID project allows us to gather measures of allocation of public funds (capture) and democratic practices. Summary statistics for all measures can be found in the appendix (E).

5.2.2 Measurement of Public Funds Allocation

To measure the allocation of public funds, we undertook a carefully designed audit in each community during Step D. Each four-person survey team included two auditors that were specifically trained to assess how much of the unconditional cash transfer was diverted by the committee, through which methods, and to whom. They visited nearby markets, interviewed beneficiaries, dealers, market users, and committee members to verify actual amounts distributed, local prices, local exchange rates, transactions, the quality of material used, the value of remaining assets and the distributional patterns of the cash transfer. After the audit, the auditors recorded the share of the \$1,000 grant that they were unable to account for.

Second, we measure citizen perceptions of embezzlement recorded through a direct question in the household survey and a list experiment. That is, during Step D, we asked ten randomly selected respondents in each RAPID village a simple direct question of whether they believed the RAPID project suffered from problems of embezzlement — like corruption

⁹We describe the field conditions in terms of team structure, infrastructure, and security conditions in the appendix (K.2).

Table 1: The RAPID Behavioral Measure

Stage	Description	Features
A	Team A schedules village meeting and conducts surveys	Initial meeting with the village chief to ask him/her to convene a public meeting at which a minimum share of the village population is required to attend. Survey is conducted among 5 randomly selected households.
	Village meeting and additional surveys	The RAPID project is described in a public village meeting. Measures of the quality of participation are taken. The village is asked to take steps towards determining how to use the RAPID funding and identify representatives (with no guidance). The population is informed that at least \$900 will be made available.
B	Collection of forms	Meeting with committee members only. Measures are taken of the village’s decisions regarding how to use funding and who is entrusted to manage it. The committee members are informed <i>in private</i> that the amount provided to villages will be \$1,000 (\$100 more than the minimum amount announced to the village).
C	Disbursement of funds by IRC and CARE	\$1,000 are disbursed in private to a select group of members identified by the management committee.
D	Auditing	Auditing activities and focus groups are undertaken to track the use of all funds, and measure capture, transparency, and accountability practices that were established.
	Follow-up surveys	Surveys are conducted among 10 randomly selected households (5 are those surveyed during Step A). Measures are included to determine the transparency of the RAPID process, the quality of participation in village decision-making, and the equity of outcomes.

Notes: Key features of the \$1,000 unconditional cash transfer program.

and nepotism — by RAPID committee members. However, since reporting sensitive answers can have reporting bias, we also elicit responses through a list experiment from these same respondents (Blair and Imai, 2012). Specifically, a randomly selected half of the respondents received a baseline list of statements, and were asked how many of them were relevant. The other respondents received the same list but with the sensitive statement of interest appended.¹⁰ Because the responses are aggregated, the respondent can have some assurance that nobody knows the answer to the sensitive question. We take the difference between the average response in the long and the short list groups as our measure for embezzlement.

Third, we also explore the inequality of benefits among households. Since most RAPID projects involved the distribution of household goods in the community, we can compute the extent to which final benefits were captured by a small elite, or instead were widely distributed. During Step D, surveyors asked the ten respondents whether their household

¹⁰The sensitive item was: “There was embezzlement in the village by the RAPID leaders (corruption, nepotism, etc.).” The baseline list included the following statements: a) The population would have liked more training by RAPID; b) The population would have liked more time to carry out the RAPID project; c) The population would have liked RAPID organizers to provide technical knowledge.

received something directly from the RAPID project. We calculate the value of each benefit and compute the standard deviation of the distributions that took place (in dollars) to represent the average difference in the amount received between two randomly selected villagers. Thus, we obtain a village level outcome that characterizes the village-level distribution of benefits.

Finally, we measure the effect on the extent to which decisions reflect the preferences of the village chiefs versus those of the villagers. To do so, we compare the predictive power of the chief’s preferences to those of a random sample of five villagers. We obtain a village level measure that indicates the degree to which the chief’s private preferences (measured during Step A before the townhall) outperform the preferences of the five panel households who were also interviewed before the RAPID project in private (also collected during Step A before the townhall), at predicting RAPID project choice (observed at Step B).¹¹

5.2.3 Measurement of Democratic Practices

We examine democratic practice along three dimensions: participation, accountability, and transparency.

Measurement of participation. First, during Step A, we count the number of villagers present in the town hall meeting of the RAPID project, and record the number of times that the average villager speaks in these meetings, as well as the dominance of men in such discussions. However, the patterns of public communication may not perfectly correspond to the actual inclusiveness of the process, for instance, if the most powerful individuals are less likely to talk in cases when town hall meetings are partly performative. We thus complement our measure of participation with the extent to which the RAPID project and the community committee that ends up being in charge of the funds are selected through participatory selection methods. That is, between Steps A and B, RAPID communities were required to select both a committee and a project as part of the terms of receiving funds. During Step B, enumerators conducted two focus groups simultaneously, one with members of the committee and a second with ordinary villagers, and coded the selection process as either electoral, through lottery, by consensus, imposed by the chief or elders, other, or unknown. We combine these four measures into one composite measure. Last, if the average villager is more likely to effectively participate, we should expect RAPID committees to have a broader representation of the population. We implemented an additional measure of participation: the composition of the RAPID committee. There was no constraint placed on

¹¹We obtain similar results if we use data on RAPID project choice from Step D. Due to attrition in the latter we have more observations from Step B.

the composition of these committees other than size (at least 2 members and no more than 8). We create a composite measure based on the number of women, the number of men, the total size, and the share of women that make up the committee.

Measurement of accountability. To measure accountability of elites, we first examine whether the community has put in place accountability mechanisms to control the actions of the committee in charge of the unconditional cash transfer funds. Was an external accountability measure (such as a distinct committee) put into place? Or was the committee required to report its actions to the community as a whole? During Step D, enumerators conducted focus groups with ordinary villagers, focus groups with at least two RAPID committee members, and interviews in private with the ten randomly selected respondents and two randomly selected committee members. We combine information from these four measures into one composite index. Next, we complement this measure of accountability by gathering information about people’s propensity to complain about the RAPID implementation. Specifically, during Step D, we asked the ten respondents in private to indicate whether or not they agreed with thirteen pre-selected complaints.¹² We create an index of the average propensity of villagers to issue complaints.

Measurement of transparency. First, we examine the degree to which the population is informed about the RAPID project. Specifically, we inform the entire community during the Step A town hall meeting that the village will receive *at least* \$900 in unconditional cash transfers. However, about a week later, the RAPID project provides \$1,000 to the RAPID village management committee. The committee thus learns about the actual amount transferred to the community in private. This introduces information asymmetry which allows us to measure the extent to which knowledge of the funding amount is subsequently shared by selected leaders to citizens. During Step D, we ask the ten randomly selected villagers to tell us the amount of the RAPID grant. We measure whether they report the correct amount. Second, we complement this measure with an incentivized behavioral measure. We randomly select two villagers in both RAPID and survey-only villages, and offer them monetary rewards to obtain their village’s school budget. We first measure the proportion of respondents who accept the task, and for those who refuse, why they do so. For those who accept and return, we compare the figures they report to the figures that we obtained as part of the direct interview with the school director. Finally, we measure the

¹²The process took too long; The organization (RAPID) did not behave well in villages; The projects selected were not the most important ones; The selected projects did not benefit a wide enough group; I had no real influence over the selection process; Disagreements were not well managed; The process was too complex; There was not enough information about the process; There was corruption (misuse of funds) in the village; The distribution of funds was not just; The project created divisions in the community; The RAPID committee was too influenced by the Chief; The RAPID committee did not represent our concerns.

quality of accounting. That is, RAPID committees were expected to keep an accounting form, and record the total amount made available for the RAPID project (out of \$1,000) and what expenditures were made. During Step D, our auditors obtain copies of these accounting documents, and measure the presence of these documents and rate their quality and consistency following a precise metric that was developed ex-ante.

5.3 Estimation of the Treatment Effect

To estimate the sample average treatment effect (Rubin, 1974) of taking part in the CDR program, we compare outcomes in Tuungane communities to outcomes in control communities, accounting for small differences in assignment propensities.

Our core specification estimates:

$$y_{ijk} = \beta_0 + \beta_1 T_j + \nu_k + \epsilon_i \quad (1)$$

where i indicates the individual, j indicates the village cluster, and k the lottery block. We include lottery block fixed effects, ν_k . The coefficient on T_j provides our estimate of the treatment effect.

In some specifications, we estimate:

$$y_{ijk} = \beta_0 + \beta_1 T_j + \beta_2 X_i + \beta_3 X_i T_j + \nu_k + \epsilon_i \quad (2)$$

This specification is used for the analysis of list experiments where the interest is in knowing whether the treatment increased the difference between long list and short list responses. In this case, the coefficient on $X_i T_j$ captures the effect of the treatment effect on the difference between long and short lists. We also use this specification for the analysis of chiefly influence, where the interest is in knowing whether the treatment is associated with a smaller difference between the influence of chief and citizen preferences.

We further clarify our strategy with respect to weighting, clustering of standard errors, controls, and multiple comparisons issues.

Weighting. Because randomization blocks sometimes contain an odd number of clusters, some randomization blocks have slightly different assignment probabilities than those with even numbers (in which case 50% of units are selected). To account for this, we estimate the mean outcome in the treatment and control groups where we weight observations using the inverse of the treatment assignment probabilities (henceforth, inverse propensity weighting). With small deviations in assignment probabilities, this has a negligible effect on our estimates

(Angrist, 1998).¹³ We report results using sampling weights in the appendix (K), which produce the same conclusions.

Controls. Controlling for randomization blocks can improve efficiency (Bruhn and McKenzie, 2009). Though not specified in our pre-analysis plan, in the main analysis, we present the results using block fixed effects. In the appendix (K.6) we show that our estimates are unchanged without them.

Clustering. The CDR program is assigned at the village cluster level, j . For those analyses that make use of individual-level data, we cluster the standard errors at the level of treatment assignment.

Multiple comparisons. For five of the fifteen outcomes, we use multiple measures and risk having multiple comparisons problems. Thus, we generate a summary of effects within each group of measures that are conceptually related, following the approach of Kling et al. (2007) and create a standardized index of measures in each group (see also Casey et al. (2013)).

6 Results

We first present the effect of participating in the CDR program on the allocation of public funds, and then on the extent to which communities use more democratic governance practices.

6.1 Impacts on Public Fund Allocation

Table 2 presents the effect of participating in the CDR program on the allocation of public funds.¹⁴ The “Control” column describes the estimated level for each measure in control communities. The subsequent column provides the size of the estimated effect of Tuungane, followed by our estimated standard error, the number of observations and the number of clusters in which data was collected.

We now describe the key results. First, the effect on the share of the \$1,000 grant that auditors were unable to account for is indistinguishable in villages that took part in the

¹³Technically we need to account for conditional heterogeneity in propensities given sampling into measurement. Thus, for example, if two units were randomly sampled from two blocks of size four and five, propensity weights would place weights of 2 on each of the four units in the first block and 2.5 on the treated units in the second block and 1.66 on the 2 control units in the second block. If, however, we sample only two of the three control units in the second block then reweighting is not needed to produce unbiased within-block estimates. Since we sample 280 out of 320 control areas into measurement we take account of this when producing propensity weights.

¹⁴See also Figure 3 in the appendix (F) for a graphical representation of these results.

CDR program from villages that had not taken part in the CDR program. On average, approximately 15% of the \$1,000 could not be verified by the teams. There is no significant difference between Tuungane and control communities. The estimated effect is very small with a small standard error. This suggests little difference in fraudulent behavior across treatment groups, although this does not itself mean that resources that could be accounted for were used well.

Second, the level of funds embezzlement in the RAPID project in villages that took part in the CDR program is indistinguishable from villages that did not. For the measure of embezzlement reported by households using the direct survey question, 15% of respondents in control communities report this to be a concern. Results are similar in Tuungane communities. When we examine instead the household reports of funds embezzlement based on the list experiment, we find again no statistically significant difference in Tuungane areas.¹⁵

Table 2: Effect on Public Fund Allocation

	Index	Control	Effect	(se)	N	Clusters
Financial Irregularities	No	0.147	-0.006	(0.02)	394	394
Embezzlement (direct)	No	0.147	-0.001	(0.018)	3623	411
Embezzlement (list experiment)	No	0.462	-0.012	(0.066)	3676	411
Inequality of (Private) Benefits	No	2.602	0.163	(0.495)	409	409
Dominance of Chief's Preferences	No	0.095	-0.019	(0.039)	2446	441

Notes: For “Embezzlement (list experiment)” and “Dominance of Chief’s Preferences” we estimate equation (2), and report estimates for β_2 in the Control column and estimates for β_3 in the Effect column, where X is the sensitive item and the chief, respectively. All analyses employ inverse propensity weights, clustering of standard errors at the level of randomization clusters, and block fixed effects. Cluster column refers to the number of unique village clusters. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Third, villages that took part in the CDR program distribute the RAPID project funds with indistinguishable levels of inequality as villages that did not take part of the CDR

¹⁵For details on the number of clusters used for different analyses see appendix K.3.

program. On average, in control communities this standard deviation is around \$2.6, and indistinguishable from Tuungane communities.¹⁶

Fourth, the preferences of the chief in villages that took part in the CDR program, and relative to the population’s preferences, are no less able to predict the RAPID project choice than in villages that did not take part of the CDR program. In control villages, the chief’s prior preferences are eight percentage points more likely than those of a randomly selected villager to coincide with actual RAPID project choice. However, the CDR program has no effect on the extent to which chief’s preferences outperform villagers’ preferences.¹⁷

In conclusion, exposure to the CDR program has no statistically significant effect on the choices made by communities during the RAPID project.

We note a final observation on project outcomes. Discussions with the project funder and implementation agencies suggested expectations that in control areas most if not all of the \$1,000 RAPID funds would be embezzled by traditional elites. Our data do not bear this out on any measure. The most pessimistic results come from the list experiment on embezzlement, where still more than half of all communities indicated no embezzlement. Other direct measures are even more positive on this front.

6.2 Democratic Practice

Table 3 presents results related to democratic practice. Before discussing individual results, we note that there is a drop in observations when moving from our measures for participation to those for accountability and transparency. The reason is that the measures for participation build on data collected during Step A and B, while our measures for accountability and transparency build on data collected during Step D. Enumerator teams were expelled from Maniema province after most of the Step A data was collected, but before much of the Step D data was collected (see appendix K.3).

Also note that our examination of democratic practices in a natural setting comes with implications for interpreting our estimand. We assess, for instance, whether the institutional intervention altered participation, which in turn may have affected choices. It is possible, however, that the intervention could have affected *willingness* to participate to prevent capture by elite but that this does not translate into greater participation because of the effectiveness of the threat. We measure effects on actual participation, not on propensity to participate.

¹⁶We note that in villages in which public goods were produced and no distributions with cash value were made this standard deviation is zero.

¹⁷The number of clusters is higher for this measure because it is based on data from Steps A and B.

6.2.1 Participation

Table 3 shows that in control communities, on average, 132 adults participated in the public meeting, one more than in Tuungane communities; a very small difference which is not statistically significant. These meetings were attended by two enumerators, who recorded patterns of social interaction. The second row shows that, on average, 15 interventions are made per meeting, with only marginally fewer interventions in Tuungane communities. Furthermore, we find that in control communities men dominate the discussion, being responsible for 70% of the interventions. The third row shows that the patterns of male dominance of social interactions are indistinguishable across treatment and control.¹⁸

Next, we examine results related to the process of RAPID project and committee selection. Approximately 43% of committees and 31% of RAPID projects were coded as selected through election, and 71% of committees and 73% of the projects were selected through either election, lottery or consensus. Table 3 shows the composite measure, which by construction averages zero in control areas and has a standard deviation equal to one. We find that there is no evidence that participation in Tuungane leads to greater adoption of participatory processes in the selection of the committee or spending plans. The estimated effect is below one tenth of a standard deviation.

Last, we examine the inclusiveness of the RAPID committee. On average, about one committee member in six was a woman (17% in control; 20% in treatment). The composite index — which includes the number of women, the number of men, the total size, and the share of women on the committee — shows that there is no statistically significant difference between Tuungane treatment and control communities.¹⁹

6.2.2 Accountability

In the majority of villages, no mechanisms had been put in place to oversee the use of RAPID funding. However, 13% of respondents indicated that an external accountability measure (such as a distinct committee) had been put into place, and another 11% indicated that the committee had been required to report its actions to the community as a whole. As the composite measure in Table 3 indicates, Tuungane did not lead to a greater propensity to put accountability mechanisms into place.

Second, we examine the propensity to complain as calculated by an index of the average propensity of villagers to issue complaints. Results in Table 3 suggest that levels of complaint

¹⁸We find similar results for dominance of the chief and elderly.

¹⁹Looking at the number of women and the share of women individually, we do find evidence that the Tuungane program had an impact.

Table 3: Effect on Democratic Practices

	Index	Control	Effect	(se)	N	Clusters
Participation						
Meeting Attendance	No	132.394	-1.199	(6.297)	455	455
Interventions in Meeting	No	14.697	-0.267	(0.483)	457	457
Dominance of Men in Discussion	No	70.255	0.161	(1.36)	442	442
Participatory Selection Methods	Yes	0.015	0.072	(0.073)	451	451
Committee Composition	Yes	0.033	0.099	(0.078)	452	452
Accountability						
Accountability Mechanisms	Yes	0.007	-0.036	(0.094)	414	413
Private Complaints	Yes	0.015	-0.01	(0.052)	3658	411
Transparency						
Knowledge of Project Amount	No	37.965	0.697	(2.384)	3699	411
Willingness to Seek Information	No	39.161	2.399	(2.335)	1407	411
Quality of Accounting	Yes	-0.026	0.011	(0.084)	399	399

Notes: Outcome measures in rows 4 to 7 and row 10 are indices and by construction have zero average in control areas. Deviations from 0 in the control column may arise because we report the weighted average of block average outcomes in control areas, which differs marginally from the overall average. All analyses employ inverse propensity weights, clustering of standard errors at the level of randomization clusters, and block fixed effects. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

are no higher in Tuungane areas than in control. Note that it could be that they are no higher because RAPID project management was better, although, as we saw above, there is little evidence to support this.²⁰

²⁰We have also conducted analysis of whether complaints are greater conditional on mismanagement. This analysis suggests a positive effect of the program. However, we note that this test was not registered nor is it identified, since mismanagement is potentially endogenous.

6.2.3 Transparency

To assess effects on transparency we first examine people’s knowledge regarding the grant amount. Table 3 shows that, on average, 38% of all respondents report the correct answer of \$1,000. However, we find no evidence that there is a difference between treatment and control communities.

Second, we examine the effect on the willingness to seek information relevant to the public. Table 3 shows that approximately 39% of those in control communities were willing to seek information (receiving one dollar for the attempt, and an additional dollar upon success). The people that refused gave various reasons: that it is not appropriate to ask for this information (76), that the respondent did not have time (75), that the exercise is strange to them (40), that the husband of the respondent refuses or would refuse the collection of this information (9), and other reasons (192). This suggests that individuals do not feel they have rights to access basic financial information. There is no significant difference between treatment and control communities.

Finally, we examine the effects on quality of accounting. We find that, on average, in 83% of the villages, the committee had their accounting form present upon arrival of the audit team during Step D. Approximately 80% of the funds were correctly accounted for as calculated by the RAPID Committee (and 80% when calculated by the audit teams). In addition, 58% of the money the committee made available for the RAPID project (of the \$1,000) was justified by receipts, and 47% was justified with receipts deemed credible by the auditing team. Table 3 presents the composite index taking these individual measures into account. We do not find evidence of an impact of Tuungane on the existence and quality of accounting.

In summary, we find no evidence that the intervention induced democratic practices.

7 Is the Null Due to Failures in Design?

Our null results might reflect a weak treatment effect, but they might also reflect weaknesses in our research design. We explore three possible reasons that could in principle produce false null results: spillover bias, differential social desirability biases, and low statistical power. In addition, we consider the possibility that results reflect a bias due to a short term elite response to the intervention (see appendix I). We also explore survey-based information that addresses the concern that the intervention itself was a poor case for finding evidence of effects, or that the intervention was poorly implemented (appendix J). Finally, we explore

concerns related to data missingness, compliance, treatment heterogeneity and specification biases (appendix K).

7.1 Spillovers

It is possible that Tuungane produced positive effects beyond treatment communities and that these positive spillovers bias our estimates of effects downwards. To address this concern we take account of indirect village exposure to treatment. We define an “ x -km indirect effect” as the effect of being within x kilometers of a Tuungane village that is part of another cluster of villages.²¹ Because of some data missingness our results assess the effects of being close to a treatment village *for which we have location data*. We ignore this detail in what follows in light of the small number of units with missing data (we have GPS locations for a total of 1,020 of the 1,120 villages). The propensity of being exposed to such a treatment effect depends not just on the random assignment of units to treatment but also on the location of any given unit with respect to others. We make use of the random assignment of Tuungane to recover these propensities, since they are determined by our original randomization. To calculate these propensities we replicate the random assignment to Tuungane to obtain 5,000 possible assignments of all units to treatment and control, employing the same scheme as used in the original randomization. We then assess, for each unit, the probability of receiving direct treatment, indirect treatment, and each combination of these. To avoid instability arising from large weights, we limit the analysis to villages that have at least a 10% to 90% probability of being in each of these groups for a given value of x . We then generate estimates of treatment effects by comparing outcomes in each combination of conditions with inverse propensity weighting using the *known* propensity for each unit of being in each condition. We test the sharp null of no effects using a randomization inference procedure (Fisher, 1935).²²

We conduct our analysis for both a 5km radius spillover treatment and a 20km radius

²¹Note that for the spillover analysis missing data affects both the set of units in the study *but also the measures of exposure to spillovers*. The strategy we use to assess the presence of spillovers is design-based in the sense that it uses information on the probability of exposure to spillovers that can be calculated from the assignment strategy. See Gerber and Green (2012), Chapter 8, for more details. We emphasize however that this does not mean that we do not depend upon a substantive model: rather we require a type of SUTVA violation such that units not depend upon the treatment condition of other units beyond those described by the “ x -km indirect exposure (Aronow and Samii, 2013).

²²That is, for each of the 5,000 re-assignments to Tuungane we calculate the estimated effect of each treatment type for each outcome of interest. Combined, these estimated effects produce a reference distribution of model root mean squared error (RMSE) under the sharp null. We compare the actual root mean squared error to this distribution and estimate how likely it is we would have obtained results as strong or stronger than the actual RMSE under the sharp null. See also: Barrios et al. (2012).

spillover treatment.²³ While our estimates of spillover effects depend on the assumption that, in each analysis, we have correctly modeled the structure of spillovers, our test of the sharp null of no effect does not (Bowers and Fredrickson, 2013).

The results (presented in appendix G) show no evidence that exposure to Tuungane had spillover effects on good governance (or direct effects, once we take account of possible spillovers). We see considerable movement in coefficients compared to our main specifications, reflecting our reduced power for these analyses, though none appear systematic. This conclusion is also supported by the following considerations. First, randomization was implemented at the level of village clusters, hence most treated villages are surrounded by treated villages and most control villages by control villages, which limits the scope for spillovers to control areas. Second, populations in control areas report low levels of knowledge about Tuungane. About 85% of the general population and 67% of chiefs in control areas reported never having heard of Tuungane.

We conclude that it is unlikely that spillover effects account for the null results we see here.

7.2 Differential Desirability Bias

Another possible bias that we consider is that control communities, expecting future conditional aid disbursement, may have managed the RAPID grant in a more democratic manner in order to please future donors. This could be the case if the RAPID project was perceived as linked to the donor community, despite our best efforts to uncouple them.²⁴ We use a

²³When we examine different conceptualizations of the treatment effect we *simultaneously alter our samples* (we illustrate this in Figure 4 in appendix G). A unit in a block with many units, but that has no neighbor within a 10km radius, has a 50% chance of receiving the direct treatment but a 0% chance of being exposed to the indirect treatment of “having a treated neighbor within 5km.” Such a relatively isolated unit would drop out of our analysis of a 5km treatment effect. The same unit however might be retained for an assessment of the effect of being within 20km of a treatment village. Villages in more clustered areas may enter the analysis set for the first analysis but not the second (since these may have a 100% chance of being indirectly treated under the first definition). In fact, analysis for the 5km (20km) radius retains 352 (298) units, with only 92 villages being in both groups. Setting x to 5 yields 516 (504) villages that are (not) directly treated, and among those 450 (570) villages that are (not) indirectly treated. Setting x to 20 yields 504 (516) villages that are (not) directly treated, and among those 874 (146) villages are (not) indirectly treated. Conditioning on these villages having a 10% to 90% probability to be in each combination retains a total of 352 villages: 49 neither direct nor indirect, 120 not direct but indirect, 74 direct and not indirect, and 109 direct and indirect. At a 20km radius these categories total, respectively, 25, 117, 41 and 115 (summing up to 298).

²⁴More specifically, the teams introduced themselves to the villages as affiliated with the Official University of Bukavu (in Maniema and South Kivu) or the University of Lubumbashi (Haut Katanga and Tanganyika) and that the RAPID project was implemented by their respective universities in cooperation with Columbia University in New York City and was funded by the British government. Although we sought to minimize any connection with IRC and CARE we also adopted a policy of no deception: if respondents asked directly about IRC or CARE involvement, team members acknowledged their involvement, emphasizing their role in

small experiment embedded in our endline survey to shed light on this possibility.

To assess directly whether villagers were motivated to please development donors, we introduced a survey variation in which we asked the ten respondents the following question: “Do you agree with the idea that elections are the best way to choose community representatives for positions with technical responsibilities?” For a randomly selected half of respondents the question was preceded by the statement “Many NGOs in the region think that elections are not the best way to choose community representatives when it comes to an appointment with technical responsibilities”; the other half was told “Many NGOs in the region are of the opinion that elections are always the best way to choose community representatives for technical posts.” Comparison of answers allows us to assess the degree to which respondents seek to provide answers that they think NGOs want to hear.²⁵

We find strong evidence for social desirability bias: in both groups individuals are 18 - 22 percentage points more likely to answer ‘yes’ following a positive prompt (full results are provided in Table 8 in appendix H). However, this bias is not affected by exposure to Tuungane. Therefore, we believe that this form of social desirability bias among control communities is unlikely to drive our result.²⁶

7.3 Weak Statistical Power

In addition, it is possible that our study finds no evidence of treatment effects, because the sample is too small to find economically relevant effects.

This question is largely answered by examination of our standard errors. To calculate the least optimistic effect that we can reject (assuming a positive effect was sought), we compute $\beta = \hat{\beta} + 1.96se$.²⁷ The results in Tables 2 and 3 for instance, suggest that in many cases we can reject small effects. For instance we can reject the null that the impact of Tuungane on financial irregularities was above 4.5 percentage points, and similarly for measures of embezzlement. Our list experimental measure of embezzlement, which relies on

disbursing funds. Moreover, the IRC and CARE International employees that visited villages to distribute the RAPID project funds during Step C were assigned to areas in which they had not worked previously so that they would not be identified as staff by populations.

²⁵More precisely, we gave both prompts to *all* respondents but randomized the order of the prompts. Though not exploited here this allows us to generate a within person measure as well as consistency bias. The results here use only the first prompt, however, which provides the cleanest results.

²⁶We note, thanks to a reviewer comment, that other types of experimenter effects might still occur. For instance, compensatory behavior in control villages could arise from the experience of not receiving funding in the past, or from greater expectations about future flows of funds. This concern, we believe, is allayed by the fact that the Tuungane program was in general not well known in control areas.

²⁷That is, we choose the β such that $\frac{\beta - \hat{\beta}}{se} = 1.96$.

interactions, is less well powered: we can only reject effects of 14.1 percentage points or more. On participation we can reject the null of 11.1 more people showing up for meetings, from a baseline of 132, or that that women account for 2.5% more of the interventions in meetings, or that treatment villages score 0.2 standard deviations higher on our participatory selection index. Our weakest evidence is on the dominance of chiefs preferences, which also depends on estimating the coefficient of an interaction effect. For this we can reject a reduction of 9 percentage points, a figure that is similar in size to the estimate of chiefly dominance in the control group.

We also conduct an ex post power calculation. Armed with information on the structure of errors in our data we can ask how large an effect needs to be to have 80% power. To answer this question we use the structure of the outcome data collected from financial irregularities and use simulation techniques to assess power under different possible effect sizes. Figure 6 in appendix (M) presents the results. With our current sample, we would be able to detect an effect size larger than 0.05. The study is thus well set up to detect treatment effects and the findings thus suggest that the program simply did not have a meaningful effect on democratic outcomes and democratic practice.

8 Conclusion

Participatory development is commonly advocated not just as a way to improve economic outcomes but as a way to transform local institutional structures. The enthusiasm often ignores the possibility that existing governance practices may have a functional purpose or are difficult to change due to culture or existing power relations. We examined the effects of an ambitious and representative participatory development program in the Democratic Republic of the Congo on subsequent democratic practice. Our findings suggest that short-run exposure to imported democratic practices does not lead to subsequent adoption of these practices.

These findings contribute to the literature on the political economy of development and to development practice in a number of ways.

First, our findings suggest that behavioral change of local elites and the general population stemming from exposure to foreign practices is not automatic. This has implications for scholarship studying long-run development and institutional change (Akyeampong et al., 2014).

Note though that our results do not suggest that local electoral processes never matter. A series of results suggests that they can. Martinez-Bravo et al. (2011) find evidence for

beneficial effects of local elections on local accountability in China; Grossman (2014) provides evidence on the effects of different types of rules. Other work finds adverse effects: Baldwin and Mvukiyehe (2015) for example examine the introduction of elections in Liberia and find evidence that they can worsen local collective action. Unlike our study however, these studies focus primarily on the effects of institutions *on the decisions made under these institutions*, rather than on the prior question of whether communities will use these institutions. In these settings persistence is often guaranteed by an external group rather than resulting from endogenous adoption. Critically too, some of these more positive results are found in settings in which institutions connect communities to larger structures – such as national governmental structures; as such they focus on settings where institutions provide a clear mechanism through which change in governance may be achieved.

Second, our findings suggest the need to rethink the strategies employed by Western aid organizations and governments to influence local institutions. Current donor-driven approaches to render decision-making more inclusive by short or medium-term interventions, which do not change power relations, may serve the population when they are implemented, but their transformative promise does not enjoy empirical support. Our results are consistent with the view that when sectors of the population are likely to have vested interests, institutional change may itself require a change in the power endowments.

More fundamentally, the logic of exogenous institutional change assumes that external action helps shift populations from one equilibrium to another equilibrium that is believed to be better. This presupposes that these populations are in a bad institutional equilibrium in the first place. This is a common view for the type of problem we are examining and certainly one shared in this case by development actors. Scholars sometimes adopt a chiefs-as-despots model and view rural institutions as captured by traditional elites. Our results suggest that both the pessimism around existing institutions and the optimism that imported institutions get picked up quickly may rest on weak grounds.

References

- Acemoglu, D., Chaves, I. N., Osafo-Kwaako, P., and Robinson, J. A. (2014a). Indirect Rule and State Weakness in Africa: Sierra Leone in Comparative Perspective. *NBER Working paper*, (20092).
- Acemoglu, D., Johnson, S., and Robinson, J. (2001). The Colonial Origins of Comparative Development: An Empirical Investigation. *American Economic Review*, 91(5):1369–1401.
- Acemoglu, D., Reed, T., and Robinson, J. A. (2014b). Chiefs: Elite Control of Civil Society and Economic Development in Sierra Leone. *Journal of Political Economy*, 122(2):319–368.
- Acemoglu, D. and Robinson, J. A. (2012). *Why Nations Fail: The Origins of Power, Prosperity and Poverty*. Crown, New York, 1st edition.
- Acemoglu, D., Ticchi, D., and Vindigni, A. (2011). Emergence and Persistence of Inefficient States. *Journal of the European Economic Association*, 9(2):177–208.
- Akyeampong, E., Bates, R., Nunn, N., and Robinson, J. (2014). *Africa’s Development in Historical Perspective*. Cambridge University Press, New York City.
- Alesina, A., Giuliano, P., and Nunn, N. (2013). On the Origins of Gender Roles: Women and the Plough. *Quarterly Journal of Economics*, 128(2):469–530.
- Angrist, J. D. (1998). Estimating the labor market impact of voluntary military service using social security data on military applicants. *Econometrica*, 66(2):249–288.
- Aoki, K. and Feldman, M. W. (1987). Toward a theory for the evolution of cultural communication: Coevolution of signal transmission and reception. *Proceedings of the National Academy of Sciences of the United States of America*, 84(20):7164–7168.
- Aronow, P. M. and Samii, C. D. (2013). Estimating Average Causal Effects Under General Interference. *Annals of Applied Statistics*, 11(4):1912–1947.
- Avdeenko, A. and Gilligan, M. J. (2015). International Interventions to Build Social Capital. *American Political Science Review*, 109(3):427–449.
- Baldwin, K. and Mvukiyehe, E. (2015). Elections and Collective Action: Evidence from Changes in Traditional Institutions in Liberia. *World Politics*, 67(04):690–725.

- Barrios, T., Diamond, R., Imbens, G. W., and Kolesár, M. (2012). Clustering, Spatial Correlations and Randomization Inference. *Journal of the American Statistical Association*, 107(498):578–591.
- Barron, P., Humphreys, M., Paler, L., and Weinstein, J. M. (2009). Community Based Reintegration in Aceh. *Indonesian Social Development Papers*, 12.
- Basri, N., Barokah, S., Monterio, C., and Beattie, K. (2004). A report on kdp mandiri: An analysis of efforts to replicate the kecamatan development program. *World Bank Technical Report 96050*.
- Beath, A., Christia, F., and Enikolopov, R. (2011). Elite Capture of Local Institutions: Evidence from a Field Experiment in Afghanistan. *Working paper*.
- Beath, A., Christia, F., and Enikolopov, R. (2013). Do Elected Councils Improve Governance? Experimental Evidence on Local Institutions in Afghanistan. *World Bank Policy Research Working Paper Series*, (6510).
- Bidner, C. and Francois, P. (2013). The Emergence of Political Accountability. *Quarterly Journal of Economics*, 128(3):1397–1448.
- Binmore, K. G. (1998). *Game Theory and the Social Contract: Just Playing*. MIT Press, Boston.
- Bjorkman, M. and Svensson, J. (2009). Power to the People: Evidence from a Randomized Field Experiment on Community-based Monitoring in Uganda. *Quarterly Journal of Economics*, 124(2):735–769.
- Blair, G. and Imai, K. (2012). Statistical analysis of list experiments. *Political Analysis*, 20(1):47–77.
- Bowers, J. and Fredrickson, M. M. (2013). Reasoning about Interference Between Units. *Political Analysis*, 21(1):97–124.
- Boyd, R. and Richerson, P. (2002). Group beneficial norms can spread rapidly in a structured population. *J Theor Biol*, 215:287–296.
- Bruhn, M. and McKenzie, D. (2009). In Pursuit of Balance: Randomization in Practice in Development Field Experiments. *American Economic Journal: Applied Economics*, 1(4):200–232.

- Cantoni, D. (2012). Adopting a New Religion: The Case of Protestantism in 16th Century Germany. Technical Report 560.
- Casey, C. (2018). Radical decentralization: Does community driven development work? *Annual Review of Economics*, 10:139–163.
- Casey, K., Glennerster, R., and Miguel, E. (2013). Reshaping Institutions: Evidence on Aid Impacts using a Preanalysis Plan. *Quarterly Journal of Economics*, 127(4):1755–1812.
- Chattopadhyay, R. and Duflo, E. (2004). Women as Policy Makers: Evidence from a Randomized Policy Experiment in India. *Econometrica*, 72(5):1409–1443.
- Chwe, M. S.-Y. (2000). Communication and Coordination in Social Networks. *Review of Economic Studies*, 67(1):1–16.
- Fearon, J. D., Humphreys, M., and Weinstein, J. M. (2009). Can Development Aid Contribute to Social Cohesion after Civil War? Evidence from a Field Experiment in Post-Conflict Liberia. *American Economic Review: Papers & Proceedings*, 99(2):287–291.
- Fisher, R. A. (1935). *The Design of Experiments*. Oliver and Boyd, London.
- Fujiwara, T. (2015). Voting technology, political responsiveness, and infant health: Evidence from brazil. *Econometrica*, 83(2):423–464.
- Gerber, A. S. and Green, D. P. (2012). *Field Experiments: Design, Analysis, and Interpretation*. W.W. Norton, New York City.
- Giuliano, P. and Nunn, N. (2017). Understanding cultural persistence and change.
- González de Lara, Y., Greif, A., and Jha, S. (2008). The Administrative Foundations of Self-Enforcing Constitutions. *American Economic Review*, 98(2):105–109.
- Greif, A. (1994). Cultural beliefs and the organization of society: A historical and theoretical reflection on collectivist and individualist societies. *Journal of Political Economy*, 102(5):912–950.
- Greif, A. and Laitin, D. D. (2004). A Theory of Endogenous Institutional Change. *American Political Science Review*, 98(4):633–652.
- Grossman, G. (2014). Do Selection Rules Affect Leader Responsiveness? Evidence from Rural Uganda. *Quarterly Journal of Political Science*, 9(1):1–44.

- Grossman, H. I. and Kim, M. (1995). Swords or Plowshares? A Theory of the Security of Claims to Property. *Journal of Political Economy*, 103(6):1275–1288.
- Guggenheim, S., Wiranto, T., Prasta, Y., and Wong, S. (2004). Indonesias kecamatan development program: A large-scale use of community development to reduce poverty. In *Scaling Up Poverty Reduction: A Global Learning Process and Conference*. Shanghai, May, pages 25–27.
- Haley, K. J. and Fessler, D. M. (2005). Nobody’s Watching? Subtle Cues Affect Generosity in an Anonymous Economic Game. *Evolution and Human Behavior*, 26(3):245–256.
- Hamilton, W. D. (1964). The Genetical Evolution of Social Behavior. *Journal of Theoretical Biology*, 7(1):1–16.
- Hartford, T. (2012). Cash Delivery on the World’s Poorest, <https://www.ft.com/content/7ce1b356-f6f9-11e1-9dff-00144feabdc0>. *Financial Times*.
- Hartford, T. (2014). One of the largest ever randomized control trials... results are in. <http://timharford.com/2012/09/one-of-the-worlds-largest-ever-randomised-trials-the-results-are-in/>. Technical report.
- Herbst, J. (2014). *States and Power in Africa*. Princeton University Press, Princeton.
- Hoffmann, K. (2014). *Ethnogovernmentality: The Making of Ethnic Territories and Subjects in Eastern Congo: Ph.D. Dissertation*. Doctoral School of Society and Globalisation, Department of Society and Globalisation, Roskilde University.
- Hoffmann, K., Vlassenroot, K., and Marchais, G. (2016). Taxation, stateness and armed groups: Public authority and resource extraction in eastern congo. *Development and Change*, 47(6):1434–1456.
- Humphreys, M., Sanchez de la Sierra, R., and Van der Windt, P. (2012). Social and Economic Impacts of Tuungane: Final Report on the Effects of a Community Driven Reconstruction Program in Eastern Democratic Republic of Congo.
- Humphreys, M., Sanchez de la Sierra, R., and Van der Windt, P. (2013). Fishing, Commitment, and Communication: A Proposal for Comprehensive Nonbinding Research Registration. *Political Analysis*, 21(1):1–20.
- King, E. and Samii, C. D. (2014). Fast-Track Institution Building in Conflict-Affected Countries? Insights from Recent Field Experiments. *World Development*, 64(1):740–760.

- Kling, J. R., Liebman, J. B., and Katz, L. F. (2007). Experimental Analysis of Neighborhood Effects. *Econometrica*, 75(1):83–119.
- Lowes, S., Nunn, N., Robinson, J. A., and Weigel, J. (2017). The evolution of culture and institutions: Evidence from the kuba kingdom. *Econometrica*, 85(4):1065–1091.
- Mamdani, M. (1996). *Citizen and Subject: Contemporary Africa and the Legacy of Late Colonialism*. Princeton University Press, Princeton.
- Mansuri, G. and Rao, V. (2013). *Localizing Development: Does Participation Work?* World Bank Policy Report.
- Martinez-Bravo, M., Padro i Miguel, G., Qian, N., and Yao, Y. (2011). Do Local Elections in Non-Democracies Increase Accountability? Evidence from Rural China. *NBER Working paper*, 16948.
- Mead, M. (1968). *New Lives for Old: Cultural Transformation -Manus, 1928-1953*.
- Mehta, J., Starmer, C., and Sugden, R. (1992). An Experimental Investigation of Focal Points in Coordination and Bargaining: Some Preliminary Results. In *Decision Making under Risk and Uncertainty*, pages 211–219. Springer.
- Newbury, D. (1991). *Kings and clans: Ijwi Island and the Lake Kivu Rift, 1780-1840*. University of Wisconsin Press.
- Nunn, N. and Qian, N. (2014). US Food Aid and Civil Conflict. *American Economic Review*, 104(6):1630–1666.
- Paluck, E. L. (2009). Reducing Intergroup Prejudice and Conflict Using the Media: A Field Experiment in Rwanda. *Journal of Personality and Social Psychology*, 96(3):574–87.
- Paluck, E. L. and Green, D. P. (2009). Deference, Dissent, and Dispute Resolution: An Experimental Intervention Using Mass Media to Change Norms and Behavior in Rwanda. *American Political Science Review*, 103(04):622–644.
- Raeymaekers, T. (2014). *Violent Capitalism and Hybrid Identity in the Eastern Congo: Power to the Margins*.
- Ross, M. L. (2001). Does Oil Hinder Democracy? *World Politics*, 53(3):325–361.
- Rubin, D. B. (1974). Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies. *Journal of Educational Psychology*, 66(5):688–701.

- Sanchez de la Sierra, R. (2017). On the origins of the state: Stationary bandits and taxation in eastern congo. *Unpublished*.
- Shayo, M. (2009). A Model of Social Identity with an Application to Political Economy: Nation, Class, and Redistribution. *American Political Science Review*, 103(02):147–174.
- Shepsle, K. A. (2006). Rational Choice Institutionalism. *The Oxford Handbook of Political Institutions*, pages 23–38.
- Sokoloff, K. L. and Engerman, S. L. (2000). History Lessons Institutions, Factor Endowments, and Paths of Development in the New World. *Journal of Economic Perspectives*, 14(3):217–232.
- Stearns, J., Verweijen, J., and Baaz, M. (2013). *The National Army and Armed Groups in the Eastern Congo: Untangling the Gordian Knot of Insecurity*. Usalama Project. Rift Valley Institute.
- Stearns, J. and Vogel, C. (2015). *The Landscape of Armed Groups in Eastern Congo*. Center on International Cooperation at New York University, New York.
- Voigtlander, N. and Voth, H.-J. (2012). Persecution perpetuated: The medieval origins of anti-semitic violence in nazi germany. *The Quarterly Journal of Economics*, 127(3):1339.
- Young, H. P. (2001). *Individual Strategy and Social Structure: An Evolutionary Theory of Institutions*. Princeton University Press, Princeton.

Appendix

A Relation to Theoretical Literature

Our study also relates to theoretical literature that sheds light on logics through which exposure to democratic practices might alter political institutions.

The sources of institutional change, and thus, how external influence might affect institutions, depend on how institutions are conceptualized. Shepsle (2006) distinguishes between two prevalent concepts of institutions (see also Greif and Laitin (2004)). In one, institutions are the rules of the game, with enforcement of those rules guaranteed outside of the game — where “rules” refers to the mapping from actions to payoffs. In the second, institutions are the equilibria of a game, which endogenously constrain behavior. In the following section, we use a simple game to describe two intervention strategies that produce observationally equivalent behavior. One strategy, Strategy *A*, affects behavior by altering expectations but without changing the primitives of the game; the other, Strategy *B*, alters behavior by altering primitives — the mapping from actions to outcomes. Both strategies yield identical outcomes and the equilibrium payoffs are the same. Equilibria shifts arising from Strategy *A* are akin to “poverty trap” arguments in the economic growth literature.

These two conceptualizations of institutions map onto two strands of research on institutional change. On the one hand, many empirical studies examine the effects of changing rules of the game alone, such as electoral rules (see Chattopadhyay and Duflo (2004) and Fujiwara (2015)). On the other hand, long-run accounts often share the view of institutions as equilibria (Boyd and Richerson, 2002). Young (2001), for example, provides an account of social institutions as patterns of behavior that may exhibit large variations across space and time without any change to primitives.²⁸ Variation in the quality of property rights regimes, norms of fairness, or tolerance for less accountable governments can also be observed across societies that share common characteristics and for which variation is potentially attributable to equilibrium selection logics (Grossman and Kim, 1995; Binmore, 1998; Young, 2001; Chwe, 2000; Bidner and Francois, 2013; Acemoglu et al., 2011). Different belief systems have been shown to support different equilibrium forms of social organization (Greif, 1994). This is similar to the view of constitutions as self-enforcing equilibria among administrators that coordinate to constrain the power of the ruler (González de Lara et al., 2008). Since in failed states, rules are often hard to change through law, many interventions seek to shift norms and practices, or expectations of behavior, aiming to induce a better equilibrium.

However, explanations of the *transition* to a new equilibrium (the theory underlying Strategy *A*) emerged from outside classical game theory. Both logics formally underpin the

²⁸Seemingly deep social structures coupled with policy choices can obtain as equilibria in environments where very different equilibria also obtain, supported by the same fundamentals (Shayo, 2009).

rationale of external interventions like the one we study here.

On the one hand, research in the lab supports the importance of leaders and moral authorities who can influence beliefs. Equilibrium-irrelevant interventions (such as labeling options or framing the context), which leaders often can change, can change the focality of equilibria, thus leading individuals to coordinate on new equilibria (Mehta et al., 1992).²⁹

On the other hand, research in the area of cultural evolution shows that while cultural norms often persist through generations (Alesina et al., 2013; Voigtlander and Voth, 2012), adoption can also occur relatively fast (Cantoni, 2012; Mead, 1968). The literature suggests that humans’ strong tendency to imitate practices perceived to be more successful can explain the evolution of culture (Boyd and Richerson, 2002). Inter-generational transmission of culture co-exists with individual optimization and the transmission of culture across individuals and across groups (Giuliano and Nunn, 2017). But, then, when do societies adopt new practices?³⁰ According to this research, the perceived success of such practices, as well as the degree of prestige of the group who practices it are important determinants of imitation and adoption (Aoki and Feldman, 1987).³¹

A.1 Institutional Logics

Consider a simple game in which two players, Strong (S) and Weak (W), can each decide in each of an infinite number of periods whether to produce using a default technology (D) or a cooperative technology (C). Say each period decision resembles a prisoner’s dilemma. If both use the cooperative technology they produce output worth 1 unit. If both stay with the default technology their yield is $d_j = .5$ for $j \in \{S, W\}$. If one uses the default technology while the other attempts to use the cooperative technology on her own the first receives

²⁹Bidner and Francois (2013) provide a more developed approach in the context of a model of accountability relations in which changes in norms occur endogenously following particular sequences of actions by leaders. Similar results would obtain from the limited rationality models in Young (2001), where expectations are based on observation of past actions and equilibria could change following a period of deviations induced exogenously.

³⁰Related applications include “social norms marketing” in Paluck and Green (2009), and Paluck (2009).

³¹The view that primitives determine institutional change has gained traction because the changes in the primitives are easier to measure, track and, unlike practices, they are easier to articulate in economic theory (Acemoglu et al., 2001; Sokoloff and Engerman, 2000). Sokoloff and Engerman (2000) in their study of the role of institutions emphasize the importance of factor endowments in determining structural inequality; indeed they highlight the “clear implication that institutions should not be presumed to be exogenous.” In the account provided by Herbst (2014), institutional variations in state capacity and responsiveness also reflect more fundamental features, notably agricultural technology and population densities. Others emphasize access to resources, such as subsoil resources (Ross, 2001) or aid (Nunn and Qian, 2014). Yet, it remains an empirical question whether culture, practices, and equilibrium selection can also play an important role in explaining the observed variation of institutional change.

free-rider yield $f_j \in (1, 2)$ while the second receives 0. In addition, players can make cash transfers to each other (assuming utility is linear in income, we treat utility as transferable).

Baseline equilibrium. With sufficient patience, the following is a subgame perfect equilibrium of this game: both players cooperate every period, each producing .5 units of value more than they would over the returns using the default technology. Player W then transfers .4 units of value to player S , and players end the round with payoffs of .6 for W and 1.4 for S . If in any period a player plays D or the appropriate transfer is not made, then all players play D in every subsequent period.

In this equilibrium S extracts 80% of what W produced over and above what she would have gained had they both played D . Following Greif and Laitin (2004) this equilibrium *is* the institution, it is sustained by equilibrium expectations of players that cooperation will only be sustained if W makes large transfers to S . In this case we might think of the political part of the institution as the 80% tax rate imposed on W .

Suppose now a third party views this equilibrium as exploitative and seeks to change outcomes. Consider two strategies they might employ.

Strategy A. The first strategy seeks to improve the lot of W by changing the equilibrium. Leaving the game intact, the third party proposes that the surplus be divided more equally, perhaps proposing that W only transfers half as much each period to R , leaving W and S with 1.2 and 0.8 respectively. The strategy is motivated by the observation that a 40% tax regime (on surplus) can also be sustained in equilibrium and so if players adopt the right expectations the new transfers will be self-enforcing. This intervention is a purely institutional intervention: it focuses on expectations and leaves the underlying game unchanged.

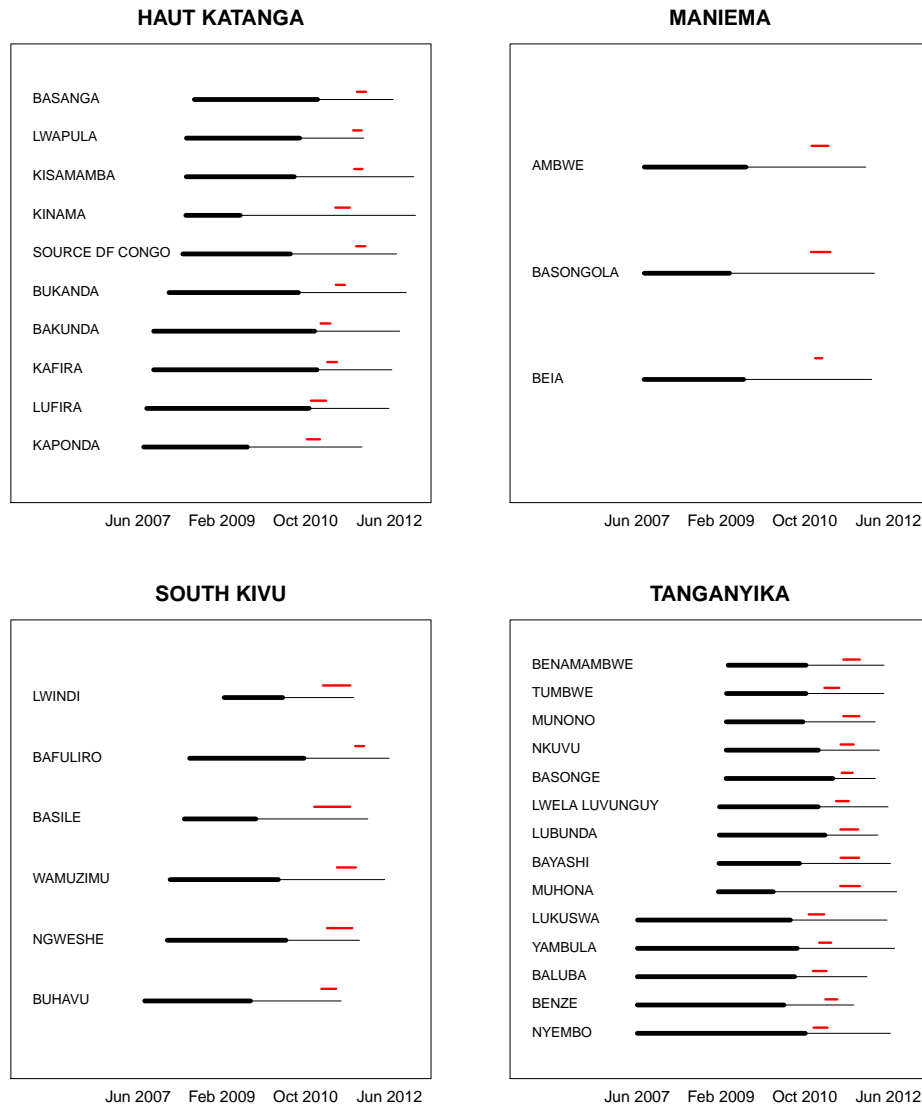
Strategy B. Consider now a second intervention in which the third party guarantees W a return of $d_W = 0.75$ instead of $d_W = 0.5$ in the event of cooperation failure. This is a structural change and has a real effect on W 's bargaining position. It means that W can now do better playing D in all periods and giving up cooperation with S . Both will still do better under some cooperative arrangement however. Say in the event of cooperation, S continued to extract 80% of W 's surplus. Then she would now force a transfer of $.25 \times .8 = .2$ and so W would be left with 0.8.

Strategy B produces the same outcome (0.8, 1.2) as achieved by Strategy A but does so without requiring a change in the approach used by the players to divide the surplus. Moreover the behaviors *on the equilibrium path* following the two interventions are the same – both players play C , each earns 1 unit and W transfers .2 units to S . The effect of Strategy B however is not due to changes in the equilibrium selected but to a change in the underlying

game (albeit one that matters only off the equilibrium path).

B Timing of Intervention and Measurement

Figure 2: Timeline of Implementation



Notes: Thin black lines indicate length of the Tuungane CDR program per chiefdom. Thick line indicates the first phase, which is the one we study here. Shorter, red lines indicate the period of measurement in that chiefdom. Source: Authors' drawing.

C Balance

We compare the different treatment conditions. Because we do not have baseline data for the villages, we make use of the data collected in 2012, where we limit ourselves to pre-treatment information (assuming no differences in recall) and variables that do not change due to the program (assuming no differences in migration). We analyze the following variables at the village level: distance (in hours) to the chiefdom capital (item QE13E in the chief survey), presence of minerals ranging from 0 to 12 indicating number of minerals present in the village (item QE20 in the chief survey), presence of infrastructure – wells, schools, clinics, churches and meeting halls – in 2006 (items CQ23-27 in the chief survey), whether previous chief was enthroned by popular choice – dummy variable for whether selected via election or plebiscite – (item CQ48 in the chief survey), and in-migration in 2006 – IDPs, returned-IDPs, refugees and repatriated refugees – (items CQ136-139 in the chief survey). At the individual level we analyze the respondents’ age (item QF9 in the household roster). Sample sizes vary due to missingness in the data. For example, few chiefs know the distance of the village to the nearest mine.

Table 4 lists the average for each variable for the treatment and control areas, and the difference between both. The d -statistic is the difference in (weighted) means expressed in (weighted) standard deviations of the control group outcome. There are no strong differences across the two groups.

Table 4: Balance

	Control	<i>Tuungane</i>	d-stat	N
Distance from major urban center	9.27	8.96	-0.02	802
Distance to village mine	5.75	6.09	0.02	470
Presence of minerals	0.84	0.78	-0.06	617
Presence of infrastructure in 2006	7.67	7.01	-0.09	722
Former chief enthroned by popular choice	0.18	0.16	-0.05	653
In-migration in 2006	8.46	8.22	-0.01	607
Age	39.77	39.3	-0.03	5411

Notes: d -statistic is the difference in (weighted) means expressed in (weighted) standard deviations of the control group outcome.

D RAPID Script Meeting Step A

We provide below the full text of the description of RAPID to communities during the general assembly meeting during Step A of the RAPID process:

“I work for RAPID and I want to talk with you about a project that we are introducing in this village. RAPID, which stands for “Research-Action through Projects for Development Impacts.” The project provides development funding from the British government and is coordinated with researchers from Columbia University in New York City and from the universities of Bukavu and Lubumbashi. The aim of the project is to provide development aid to your community while at the same time contributing to scientific research to better understand your priorities and needs.

Your village and other villages were selected in a lottery involving all the villages in this territory for the program. The program will provide a grant of at least \$900 (perhaps more) in international funding to implement a quick impact project. In this project we will let the community decide how best to use the funds.

Your chief [name] gave us permission to hold this meeting as a prerequisite for participation in the project. The aims of this meeting are to inform you of the program, to provide you the opportunity as members of the village to ask us any questions about the project, and to offer a forum for discussion on development priorities in this village and use of these funds.

There are a few requirements for participation in this project, and it is important to us that you understand them:

1. First, we want the community to decide how to use the project funds. Following this meeting, your village will have seven days to decide how to use the funds. The total funding guaranteed for this community is at least \$900. It is up to you as a village to decide the best use of funds. There are no restrictions on the use of funds, except they must be used to benefit the community and be spent out by you in the next 50 days. For this reason we encourage you to use the funds to assist members of the community through projects such as purchasing and distributing seeds, tools, large participatory work or other projects that support the well-being of this community. These funds may also be distributed to community members to use at their discretion. We prohibit the use of these funds to purchase any item whose purpose is to harm others.
2. Second, we are asking the community to identify people to represent the village for this project. These individuals will be responsible for carrying out the accounting of the use of these funds. It is up to the community to decide who these people will be over the next seven days. You are free to choose any person or persons that you feel are most appropriate to act as representatives.
3. Third, we ask you to complete this form [show form BP1] to return it on [date]. It is the Project Description Form. I will leave it with you today to complete over the next six days. The information in the form will contain the decisions you have made for the

project. A representative of Project RAPID will return in six days to collect this form. We will not be able to make the grant payment if you do not complete this form.

4. Fourth, among the questions I ask you to fill out on the form are: who are the individuals who will be responsible for managing these funds?; which project the community has chosen?; and what is the budget of such a provisional project?
5. Fifth, we ask that in two months, representatives of the community for the project RAPID provide us with an accurate accounting of the usage of funds, with evidence. This is to facilitate our understanding of the priorities of your village, as part of our research.
6. Finally, in accepting this project you also accept that the use of Project RAPID funds will be subject to an audit. What will this look like? We will send teams to implement an audit in certain villages participating in the program: if this village is audited, we will examine what the village has done with project funds. The findings will contribute to our study of the needs of eastern Congo.

Information on the disbursement of funds will be provided when collecting Project Description Forms from the representatives chosen by the community for the management of funds. Following receipt of these funds, your village will spend out these funds for your chosen project over the next 7 weeks (49 days), as is compatible with the project.

Do you have questions about this process? Would you like to participate in this project?

As we said before, there is a research component linked with this project. It is important for us that you have a good understanding of what is involved in this research so that you can use that understanding either agree or refuse to take part in it. As this project is implemented we will seek to hold a series of interviews with members of this community. These interviews will all be anonymous interviews. The aim of these is to understand the community's priorities. It is important that you understand that if you choose to be interviewed your responses will be kept anonymous.

Another part of our research will be on decision making during community meetings. Collecting measures during discussions helps us to understand more about this community and its priorities. Again we will only do this if the community agrees to this and in all cases information that is recorded will be done in a way that conserves anonymity.

Before asking for your consent we want to note that this research does not bring risks, but nor does it bring direct benefits for you. By improving our understanding of community priorities in eastern Congo this research seeks to contribute to an improvement in the quality of development aid throughout the area.

Do you consent to us collecting this data to help with this research?"

E Summary Statistics

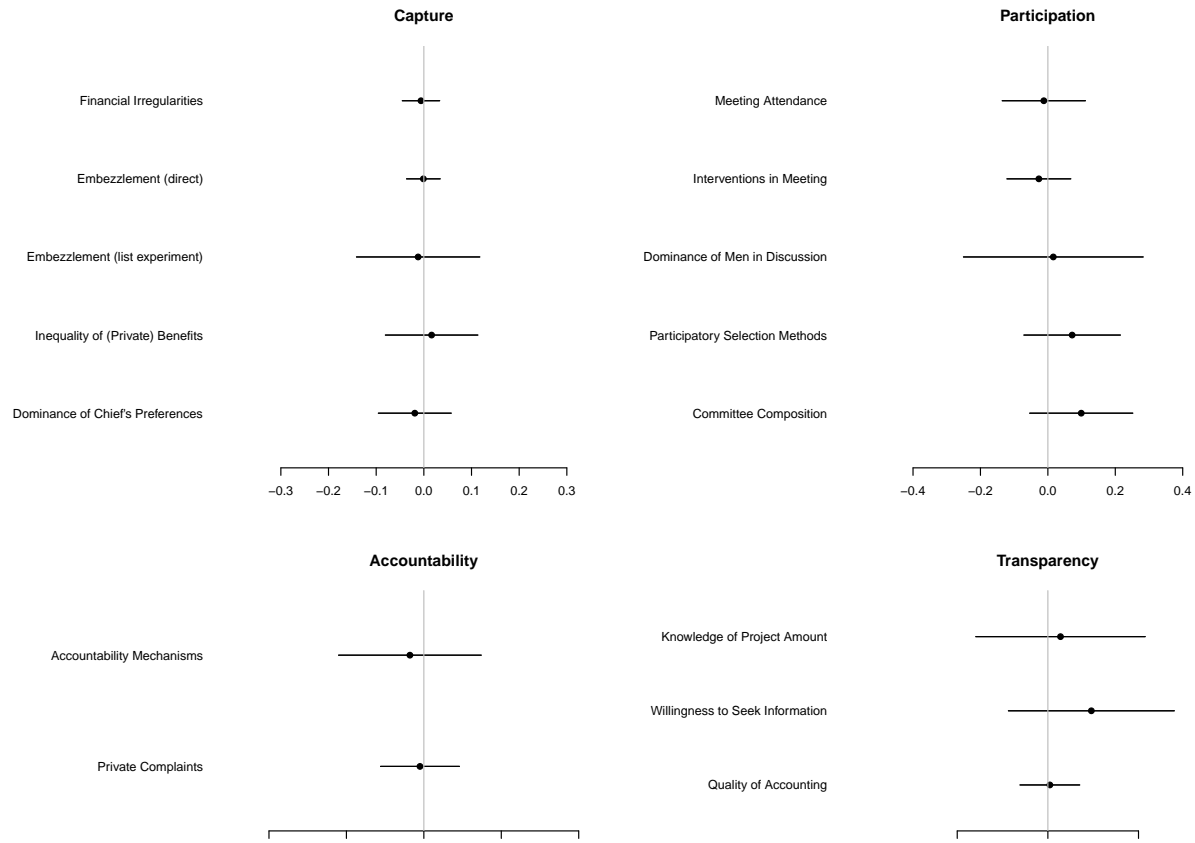
Table 5: Summary Statistics

Variable	RAPID only	Step	Index	N	Mean	SD	Min.	Max.
RAPID	NA	NA	NA	1120	0.5	0.5	0	1
TUUNGANE	NA	NA	NA	1120	0.5	0.5	0	1
Financial Irregularities	Yes	D	No	394	0.15	0.21	0	1
Embezzlement (direct)	Yes	D	No	412	0.15	0.22	0	1
Embezzlement (list experiment)	Yes	D	No	408	0.46	0.6	-1.21	2.79
Inequality of (Private) Benefits	Yes	D	No	409	2.74	5.89	0	35.84
Dominance of Chief's Preferences	Yes	A,B	No	435	0.09	0.48	-0.26	0.74
Meeting Attendance	Yes	A	No	455	131.13	79.68	20	508
Interventions in Meeting	Yes	A	No	457	14.54	5.52	1	60
Dominance of Men in Discussion	Yes	A	No	442	70.71	15.02	0	100
Participatory Selection Methods	Yes	B	Yes	451	0.06	0.99	-1.49	1.24
Committee Composition	Yes	B	Yes	452	0.07	0.96	-2.67	2.07
Accountability Mechanisms	Yes	D	Yes	414	-0.01	1.02	-2.11	3.03
Private Complaints	Yes	D	Yes	412	0.02	0.68	-0.85	2.2
Knowledge of Project Amount	Yes	D	No	411	38.48	28.07	0	100
Willingness to Seek Information	No	D	No	779	38.81	39.48	0	100
Quality of Accounting	Yes	D	Yes	399	-0.03	1.06	-3.4	1.58

Notes: Summary statistics given at the village mean level. The RAPID column indicates whether data was available only in locations in which the RAPID project was introduced. The mean effects column indicates whether a mean effects index was used in place of multiple related outcome variables. For variables analyzed as interactions we report here the difference in mean outcomes between the two relevant groups (ie for the list experiment the difference between the mean outcomes for the long list and short list subjects, and for chief dominance between the chiefs outcome and average citizen outcome).

F Graphical Representation of Main Results

Figure 3: Allocation of Public Funds and Democratic Practices

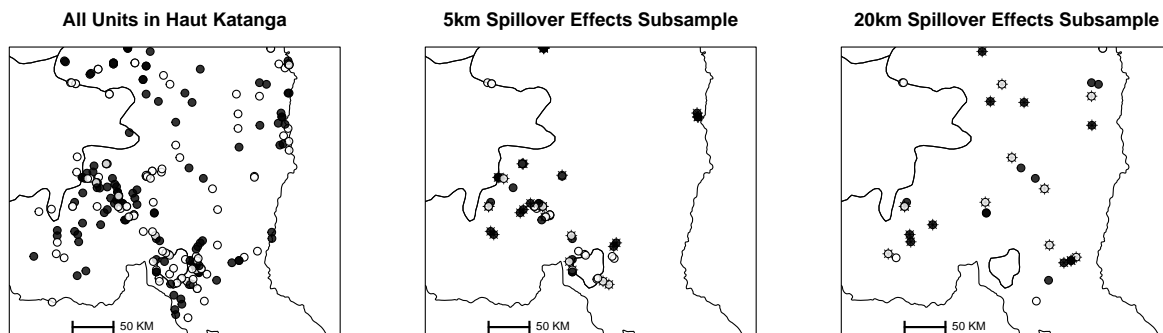


Notes: Treatment effects reported by family. All analyses employ inverse propensity weights and block fixed effects. Individual-level analyses cluster the standard errors at the level of randomization clusters. The figure does not show estimates reported in units of standard deviations of the outcome in the control group. Estimates for inequality of benefits, meeting attendance, interventions in meetings, dominance of man in discussion, knowledge of project amount, and willingness to seek information are rescaled by a factor of 1/10 to make comparisons across all outcomes easier.

G Spillovers

Figure 4 gives an illustration of areas that are included and excluded from the spillover analysis. Units are included only when they have a non-zero probability of being in all combinations of direct and indirect treatment conditions.

Figure 4: Population for Assessment of Spillover Effects



Notes: First panel shows the distribution of all treatment and control villages in a section of Haut Katanga. The middle panel shows the sub-sample of villages that had moderate ($0.01 - 0.9$) propensities of being exposed to direct *and* indirect effects of treatment, using a 5km rule for indirect exposure. Indirectly treated units are marked with a cross (and these may themselves be directly treated or not). The right panel shows the corresponding subset for a 20km rule. Changing the definition of the spillover treatment *changes the subset of cases* that have a non-extreme propensity of being exposed to spillovers. Source: Authors' drawing.

Tables 6 and 7 give the results of the spillover analysis that uses inverse propensity weights and randomization inference.

Table 6: Spillovers at 5km

	Direct	Indirect	RMSE	(<i>p</i>)	N
Spillovers at 5km					
Financial Irregularities	0.07	-0.04	0.43	0.94	156
Embezzlement (direct)	0.11	-0.04	0.42	0.83	163
Embezzlement (list experiment)	0.1	0.1	1.27	0.52	163
Inequality of (Private) Benefits	0.89	0.41	8.22	0.22	163
Dominance of Chief's Preferences	-0.05	0.09	0.79	0.5	157
Participation					
Meeting Attendance	-14.65	1.07	138.85	0.77	171
Interventions in Meeting	0.28	0.45	9.45	0.4	172
Dominance of Men in Discussion	1.17	0.23	29.41	0.8	169
Participatory Selection Methods	0.26	0.16	2	0.73	170
Committee Composition	0.23	-0.16	1.85	0.54	170
Accountability					
Accountability Mechanisms	-0.02	-0.06	1.95	0.81	164
Private Complaints	0.34	-0.03	1.41	0.89	163
Transparency					
Knowledge of Project Amount	-2.28	2.63	52.02	0.92	163
Willingness to Seek Information	7.1	-6.43	81.6	0.94	301
Quality of Accounting	-0.22	-0.13	2.03	0.59	157

Notes: Spillover effects estimated using a regression model of the form $Y = \alpha Direct + \beta Indirect + \gamma Direct \times Indirect$ where both the direct and indirect measures are normalized to have zero means. Average direct and indirect effects are then given by α and β . RMSE is used as a test statistic for the randomization inference and the p value reports the probability of such a low RMSE under the sharp null of no effects.

Table 7: Spillovers at 20km

	Direct	Indirect	RMSE	(<i>p</i>)	N
Spillovers at 20km					
Financial Irregularities	-0.02	0.01	0.29	0.24	119
Embezzlement (direct)	0.03	0.03	0.32	0.54	126
Embezzlement (list experiment)	-0.02	-0.16	1.01	0.49	126
Inequality of (Private) Benefits	-0.16	-0.55	5.99	0.89	126
Dominance of Chief's Preferences	0.12	-0.03	0.74	0.57	120
Participation					
Meeting Attendance	6.53	-42.05	146.56	0.98	141
Interventions in Meeting	-0.22	0.79	9.54	0.61	141
Dominance of Men in Discussion	0.28	0.84	29.67	0.91	133
Participatory Selection Methods	-0.12	-0.01	1.69	0.73	141
Committee Composition	0.41	0.34	1.48	0.53	142
Accountability					
Accountability Mechanisms	-0.1	-0.42	1.77	0.49	126
Private Complaints	0.02	0	0.84	0.5	126
Transparency					
Knowledge of Project Amount	-4.46	-1.3	41.61	0.38	126
Willingness to Seek Information	-2.14	11.36	68.11	0.2	241
Quality of Accounting	0.06	-0.14	1.33	0.3	120

Notes: Spillover effects estimated using a regression model of the form $Y = \alpha Direct + \beta Indirect + \gamma Direct \times Indirect$ where both the direct and indirect measures are normalized to have zero means. Average direct and indirect effects are then given by α and β . RMSE is used as a test statistic for the randomization inference and the p value reports the probability of such a low RMSE under the sharp null of no effects.

H Differential Desirability

This section presents the social desirability test. Table 8 presents the difference in the propensity to respond yes to the question "do you agree with the view that elections are the best way to choose community representatives to serve in positions that requires technical expertise?" as a function of the prompt that was given prior to the question. The table shows that, while the prompt has a significant effect on the proportion of individuals who answer yes – indicating social desirability bias – such effect is indistinguishable in treatment and control.

Table 8: Social Desirability Test

	Positive prompt	Negative prompt	Difference	(se)
Control	0.642	0.843	0.201	0.02
<i>Tuungane</i>	0.65	0.859	0.21	0.021
Difference	0.007	0.016	0.009	
(se)	0.025	0.019	0.029	

Notes: N=3,802. Share of individuals answering 'yes' to the question "Do you agree with the view that elections are the best way to choose community representatives to serve in positions that require technical expertise?" * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

I Elite Backlash Against Loss in Power

Since village chiefs were actively excluded from the Tuungane program, they might have had incentives to seek compensation during RAPID. In this case the null result could reflect unusually strong incentives for traditional leaders to engage in capture in treatment groups, coupled with strong restraints induced by bottom up pressures following the intervention. In Section 6 we found that the implemented RAPID projects (obtained during Step D) coincide better with the stated preferences (taken during Step A before the village meeting) of the chief than those of the villagers. We interpreted this as possible evidence for chief dominance. To explore whether the chief captured the RAPID process, and especially so in Tuungane areas, we investigate whether in Tuungane areas members of the RAPID committee are more closely related to the village chief.

To measure network proximity, we collected detailed friendship and kinship data among randomly selected villagers, which includes their relationship with all committee members as well as with the chief. We also collected detailed friendship and kinship data among all committee members, which includes their relationship with the village chief. We then create a measure of family connection to the committee using the Hamilton index.³² We find that neither the population nor the chief are closely related to RAPID committee members. The average score on our index for the population is 3.49%, while the score for the village chief is 4.45%. This difference is statistically significant, and amounts to the chief adding a first cousin to the RAPID committee.³³ We find no difference in this kinship proximity between Tuungane and non-Tuungane areas, however.

Other measures also confirm that the chief did not disproportionately dominate RAPID procedures in Tuungane areas. During Step B, our enumeration team led focus groups with ordinary villagers to learn whether the process of RAPID committee and RAPID project selection was electoral, by lottery, by consensus, imposed by the chief, by elders, other or unknown. Very few people (around 5%) find that the chief imposed RAPID project selection or committee member selection and equally so in treatment and control areas. Finally, during Step D we directly ask individuals whether the RAPID committee was controlled by the chief. Around 26% of the 2,514 individuals answer in the affirmative. However, from a menu of thirteen complaints less than 5% of the respondent find this to be the most important complaint. Moreover, there are no differences in reporting across treatment and control

³²The Hamilton index measures the biologic relatedness between two individuals: for a parent-offspring or full sibling relationship this index is 50%, for an aunt/uncle or nephew/niece relationship this is 25%, etc. See: Hamilton (1964). Applied to the group, if for example, two members of the RAPID committee, out of the five, are children of the chief and one is a nephew, the chief's Hamilton score is 25%.

³³Note, however that in almost 63% of the villages have no relationship at all to the chief.

areas. We thus conclude that the null result reported in this paper does not reflect chiefs' response to Tuungane.

J A Bad Case?

One possible explanation for weak evidence of effects is that this was simply a weak intervention: it was poorly implemented or not typical of the kind that researchers or policy makers expect to generate strong effects.

We have direct evidence, however, on the extent to which development funders and implementers supporting this CDR program expected that it would produce strong effects. To find out, and prior to launching our endline data collection, we ran a small survey with the population of regional program implementers and CDR program directors (12 respondents) as well as a (convenience) sample of seven researchers working in eastern Congo and Rwanda on related issues. The survey simply elicited beliefs regarding likely impacts on each of the outcomes in different categories. It was not incentivized. The responses showed variation from item to item – which suggests that respondents were not simple optimists. Two thirds of program implementation respondents reported that they thought it “improbable” that beneficiaries would allocate more time to income generating activities; none thought it very likely that household incomes would increase. Yet, all but one thought it possible or very likely that there would be improvements in each of three distinct dimensions of governance outcomes. Half thought it very likely that villages would manage CDR program in a more transparent and equitable way. Researchers were more optimistic about effects on participation, but considerably more skeptical that traditional leaders would become more accountable (most researchers reported that they would not).

Access to this information is valuable for the simple reason that it was formed prior to data gathering. If the weakness of the intervention seems obvious after the results are in, our information on priors supports the idea that the lessons may extend nevertheless to cases that are currently believed to be models. Overall, prior beliefs reflect confidence that CDR is an effective model.

K Robustness

In the text we discuss concerns related to spillovers and to social desirability biases. Here we describe issues related to attrition and data missingness, noncompliance, treatment heterogeneity, and specification sensitivity.

K.1 Sampling weights

The main analyses presented focus on sample average treatment effects. When we collect household-level data, we sample ten households in each village. Failing to account for heterogeneous sampling probabilities would result in a biased estimate of the average household in the population of interest. Similarly, in each household, we interview one respondent randomly selected within the household. Similarly, failing to account for heterogeneous sampling probabilities within households would produce a biased estimate of the population average. In Table 10, we thus employ sampling weights, derived from the sampling procedure within village across households, and within households, to account for differences in sampling probabilities across individuals in the population. These obtain the population average treatment effect (as opposed to the sampling average treatment effect in the main result). The results are unchanged.

K.2 Challenges to Implementation

The data collection effort was a very large undertaking implemented by almost 100 surveyors and their corresponding supervision and management structure over the course of over a year in a region the size of France but without any of the infrastructure. We provide a brief account of the logistics of data undertaking.

Research Teams. Multiple teams were engaged in implementing RAPID and gathering outcome data. Each province had two teams for step A, which each consisted of a RAPID project facilitator and an assistant. A teams were responsible for introducing the RAPID project to the village chief and to the village during a general assembly and for conducting a set of surveys (to be discussed in more detail below). One B team in each area visited the villages a week after Team A. These teams were responsible for meeting the committee and conducted focus groups to learn how both the committee and the RAPID project was chosen. These teams included the Provincial Supervisor who had a satellite phone in order to call the IRC or CARE International headquarters, so that a C team could visit the village to distribute the RAPID project funds. During Step C, disbursement was done by IRC or

CARE staff but without identifying themselves as such. Approximately 48 days later, after the implementation of the RAPID project, both RAPID and non-RAPID villages were visited by D Teams. D teams in RAPID villages included three enumerators and one auditor. The latter did a detailed investigation into how RAPID grants were spent and, where applicable, located beneficiary populations. The survey-only villages consisted of only two enumerators. In addition to A, B, C and D Teams each province had two Super-assistants: one responsible for Step A to Step C, and one for Step D. Super-assistants visited teams to collect and backup data, photos and GPS coordinates, and ensured quality control. These staff were hired and directed by leads at the Universities of Bukavu and Lubumbashi. Finally, there were two Regional Evaluation Coordinators, hired by the IRC: one based in Bukavu (South Kivu) and one in Lubumbashi (Haut Katanga). They were responsible for supervision of implementation, and monitoring the data collection and its quality. These Coordinators were in daily contact with the Columbia University research team and worked closely together with the Research and Evaluation Coordinator of the IRC. Between June and December 2010 two of the authors were based in Congo to launch the RAPID project.

Security. The area where the research took place is marked by high levels of insecurity, especially in South Kivu. The security of the teams was a major concern throughout and teams were not allowed to visit a village before receiving security clearance from the IRC's security team. The latter had contact with the major actors such as the United Nations peacekeeping forces, the DRC government and others. Despite the precautions undertaken we did encounter some security issues: 31 villages were not visited due to security risks; one team was ambushed and had to hand over their equipment; and one IRC staff member was abducted (and subsequently released unharmed) during the implementation of Step C. In particularly risky areas of South Kivu, Step C was undertaken through accounts in local credit offices (COOPECs) rather than having cash delivered by a field agent.

The challenges to implement a data collection exercise in this area account for most of the missing responses, which we describe in detail next.

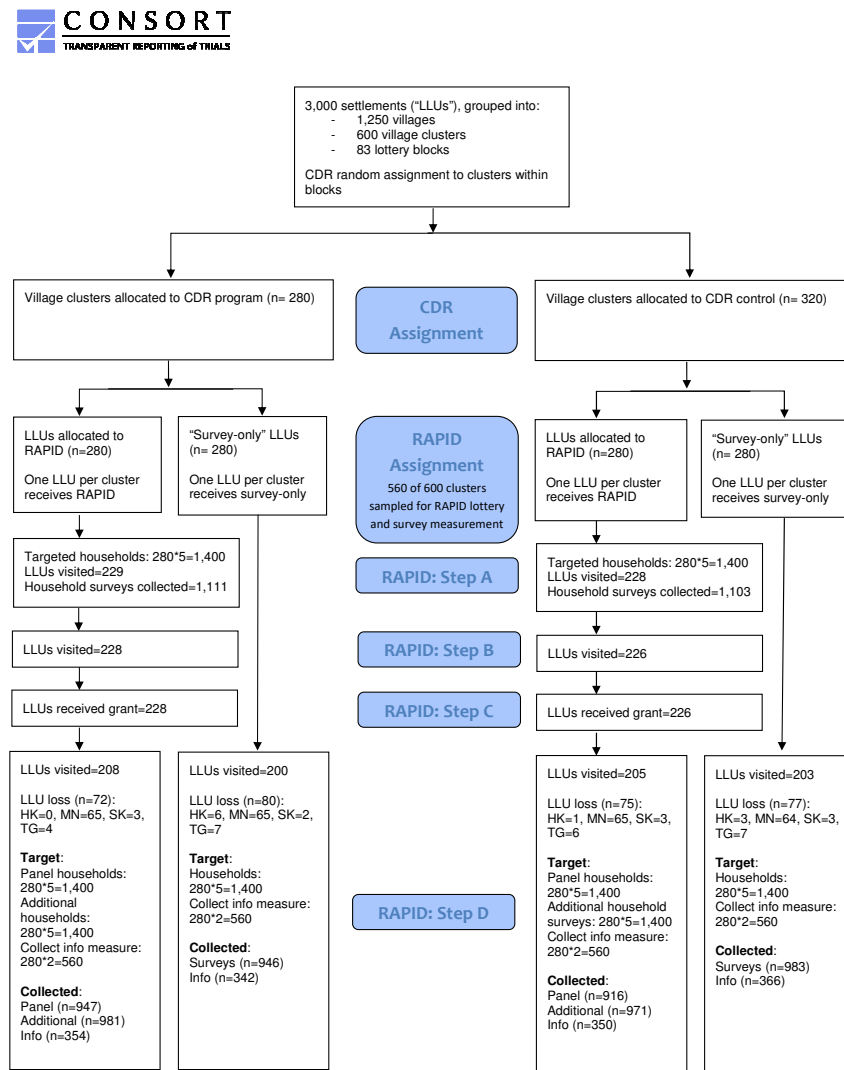
K.3 Attrition and Missing Responses

A first threat to validity stems from missing responses. Our study involves a complex collection of units – individuals, grouped into settlements, grouped into villages, grouped into village clusters, grouped into lottery blocks. Attrition in this project took place at different levels, with different implications. The most important reason for attrition, which took place due to expulsion of our teams from the Maniema province, is at the block level arising from political instability. This attrition is by design balanced for treatment and

control groups. Other attrition takes place at the level of individual villages or individual respondents.

We describe various threats arising from attrition in turn.

Figure 5: Diagram Summarizing the Organization of Units, Assignment and Measurement Strategies, and Sources of Attrition. Based on CONSORT flow chart. Abbreviations: LLUs = Lowest Level Units (settlements); HK = Haut Katanga, MN = Maniema, SK = Sud Kivu, TG = Tanganyika.



The study was designed to gather data in a sample of 1,120 villages, half of which were randomly selected for the RAPID project. Different targets were set for different items but the most common data (the household survey) were to be gathered for ten households in RAPID villages and five households in survey-only villages. Given that there were 560

RAPID villages and 560 non RAPID villages this makes a total of 8,400 households (for some items gathered only in RAPID or only in survey-only areas, the targets were 2,800). However, the survey teams successfully collected final (Step D) in only 816 villages (413 RAPID and 403 survey-only villages) and from 5,744 households. The full complement of targeted data was not gathered for a number of reasons.

Figure 5 presents a CONSORT-style flow chart with the details on the number of villages targeted and visited, and surveys targeted and collected.

The figure shows that the most significant site of missing data is Maniema province. Political tensions in the run up to the November 2012 presidential elections led to the expulsion of the Maniema teams shortly after the launch of Step D. This led to the loss of 130 RAPID villages and 129 survey-only villages for all measures based on Step D, or involving a combination of steps in this region (the data loss was greater for Step D than for Step A and Step B data, which were more advanced at the time of the expulsion). This loss covered entire lottery blocks, affecting treatment and control units alike. While it affects the range of areas to which our results can speak, as well as our statistical power, this loss is not related to the treatment status of units and is thus unlikely to induce bias.

Other sources of missing data were the inaccessibility of some regions for safety and security reasons; failures in the field, ranging from loss, damage, or theft of tablets, water damage to paper surveys, or enumerator error in the implementation of surveys or particular questions; survey non-response; and non-response on particular questions by subjects. Figure 5 shows that there is balance between Tuungane and control areas, and between RAPID and survey-only villages, for both village-level and survey-level attrition.

K.4 Noncompliance

A second threat to validity is that some areas that were selected by lottery to participate in Tuungane did not participate, and vice-versa. Survey data indicates that approximately one in seven chiefs either deny that Tuungane took place in a Tuungane community, or claimed that it did take place when according to records it did not. For all cases with discrepancies between our data and chief reports we asked the IRC to confirm whether the CDR program did or did not take place in these areas. IRC records of where Tuungane did take place matched our records of where Tuungane ought to have taken place in 77% of these ambiguous cases. This suggests that the discrepancy is due either to weak impact, poor recall by chiefs, or enumeration error. The check left 51 cases out of 806 of possible noncompliance and/or database error. For the analysis in this paper we use our database measure of units

selected by lottery which, assuming our database is correct, can be interpreted as “intent to treat” effects (albeit with a high compliance rate). In a robustness test we analyze results under the assumptions that our database is incorrect, that the IRC data is correct, and there is no failure of compliance. Our results (see Table 9) are similar.

K.5 Treatment Heterogeneity

As seen in Figure 2 there is heterogeneity both in the timing and length of CDR program implementation and the timing and length of data collection relative to CDR program implementation. Broadly the research schedule sought to follow the timing of the start date of implementation of Tuungane in each area, although the research schedule was more compressed. While the timing of CDR program initiation spanned approximately two years (with the first lottery date being in July 2007 and the last in April 2009), the data gathering spanned approximately one year (with the first village that was visited with step A of RAPID in October 2010 and the last villages visited for step A in October 2011). Thus, in general, and by design of the research, areas that launched late also had a shorter lag between start and measurement. The median gap was 1,185 days, and 90% of cases had a gap between 871 and 1,202 days. These timing decisions however all took place at the level of lottery blocks, all units in lottery blocks were first exposed to the CDR program at the same time (although CDR programs started at different times) and were visited by the research team at the same time, thus ensuring strong balance in timing issues between treatment and control areas at the block level. The implication of this heterogeneity is that the results should be seen as the average of a set of experiments that varied in time to measurement.

K.6 Specification

We also undertake a series of robustness tests to examine the extent to which the non results are sensitive to various features of our specification. First, we estimate all effects at the village level, where the variables are aggregated using individual sampling weights. The village level analysis is then done using propensity weights only, limiting the extent to which extreme sampling weights can influence cross village comparisons. Second, we generate results (at the village level) using propensity weights adjusted to assess village level sample average treatment effects rather than population average treatment effects. These weights have lower variance and may provide more precise estimates. Our results (see Table 9) are robust to these different specifications.

Table 9: Robustness

	Base	Alt. Treat.	Alt. Specifications		
	(se)	(se)	Village (weighted)	No block FEs	Village (unweighted)
	(1)	(2)	(3)	(4)	(5)
Financial Irregularities	-0.006 (0.02)	0.001 (0.02)	-0.004 (0.021)	-0.004 (0.021)	-0.004 (0.021)
Embezzlement (direct)	-0.001 (0.018)	-0.005 (0.019)	0.006 (0.01)	-0.001 (0.021)	0.006 (0.022)
Embezzlement (list experiment)	-0.007 (0.054)	-0.017 (0.055)	0.031 (0.023)	0.013 (0.06)	0.031 (0.022)
Inequality of (Private) Benefits	0.163 (0.495)	0.182 (0.49)	0.206 (0.588)	0.206 (0.588)	0.206 (0.588)
Dominance of Chief's Preferences	0.024 (0.03)	0.034 (0.03)	0 (0)	0.017 (0.031)	0.008 (0.005)
Meeting Attendance	-1.199 (6.297)	-1.482 (6.347)	-1.983 (7.403)	-1.983 (7.403)	-1.983 (7.403)
Interventions in Meeting	-0.267 (0.483)	-0.12 (0.494)	-0.391 (0.508)	-0.391 (0.508)	-0.391 (0.508)
Dominance of Men in Discussion	0.161 (1.36)	-0.026 (1.375)	0.158 (1.49)	0.158 (1.49)	0.158 (1.49)
Participatory Selection Methods	0.072 (0.073)	0.08 (0.073)	0.072 (0.094)	0.072 (0.094)	0.072 (0.094)
Committee Composition	0.099 (0.078)	0.078 (0.079)	0.083 (0.096)	0.083 (0.096)	0.083 (0.096)
Accountability Mechanisms	-0.036 (0.094)	-0.017 (0.096)	-0.029 (0.1)	-0.029 (0.1)	-0.029 (0.1)
Private Complaints	-0.01 (0.052)	-0.015 (0.054)	0.014 (0.028)	-0.006 (0.066)	0.011 (0.067)
Knowledge of Project Amount	0.697 (2.384)	0.06 (2.411)	0.646 (1.647)	0.414 (2.837)	1.436 (2.86)
Willingness to Seek Information	2.399 (2.335)	1.579 (2.342)	1.964 (2.289)	2.846 (2.967)	3.661 (3.004)
Quality of Accounting	0.011 (0.084)	-0.01 (0.086)	0.013 (0.105)	0.013 (0.105)	0.013 (0.105)

Notes: ‘Base’ corresponds to the results reported in Table 2. “Alt. Treat.” are results using a treatment variable that uses IRC’s classification of treatment in cases in which databases disagreed. “Village (weighted)” are results in which all variables are aggregated to the village level using individual sampling weights. Estimation includes inverse propensity weights. “No block FEs” are results specified in Table 2 but without block fixed effects. “Village (unweighted)” are results at the village level aggregated without sampling weights, weighted by inverse propensity weights. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

K.7 Heterogeneous Effects

Finally, our results may mask positive effects for population sub-groups. In particular, it is plausible that democratization was already advanced in most areas, and only in areas subject to a lot of capture by local elites did the program have an effect. We rule out this alternative explanation by identifying pre-existing levels of capture and estimating the heterogeneous effects of Tuungane by this pre-treatment characteristic.

To measure pre-treatment capture, we use three different indicators, based on data from

our surveys with village chiefs. First, we construct indicators of community mobilization. We identify villages without village association or committee before the start of Tuungane; this leaves us with 170 of the 358 villages. Second, we identify villages that had no classrooms in July 2006. This is the case for 194 of the 358 RAPID villages. We believe that areas where capture is effective will have lower public goods provision, especially for basic services such as education. Third, we look for the degree of competition by which current chiefs acquired their position. Absence of competition to local chiefs has been described to be a major driver of chiefs' capture of communities and civil society (Acemoglu et al., 2014b). To measure the degree of competition, we identify villages where the village chief inherited his position from his father. While 37% of chiefs' positions were inherited, 10% were chosen by elders, 25% were chosen by the local Mwami (traditional head of a large territory), 14% were chosen by other chiefs and 14% were chosen by elections.

Table 10 presents the results from the subgroup analysis. We find very little evidence of heterogeneous effects in favor of positive effects among captured communities. In addition to a few positives, an equal number of coefficients are negative.

Table 10: Heterogeneous effects of CDR program exposure, by initial institutions

	Base (se)	Projects (se)	Committees (se)	Inherited (se)
Financial Irregularities	-0.006 (0.02)	0.006 (0.031)	-0.047 (0.035)	0.004 (0.042)
Embezzlement (direct)	-0.001 (0.018)	-0.009 (0.03)	-0.02 (0.029)	-0.004 (0.035)
Embezzlement (list experiment)	-0.007 (0.054)	-0.046 (0.092)	-0.092 (0.097)	-0.052 (0.103)
Inequality of (Private) Benefits	0.163 (0.495)	1.526 (0.961)	1.596 (0.855)	-0.379 (1.139)
Dominance of Chief's Preferences	0.024 (0.03)	-0.032 (0.047)	0.007 (0.05)	-0.036 (0.063)
Meeting Attendance	-1.199 (6.297)	-4.547 (10.954)	-1.248 (13.175)	-9.803 (15.321)
Interventions in Meeting	-0.267 (0.483)	-1.046 (0.939)	-0.358 (0.903)	-1.233 (1.125)
Dominance of Men in Discussion	0.161 (1.36)	-2.196 (2.566)	-1.131 (2.433)	0.749 (2.956)
Participatory Selection Methods	0.072 (0.073)	0.328 (0.157)	0.389 (0.163)	0.106 (0.169)
Committee Composition	0.099 (0.078)	0.091 (0.166)	-0.006 (0.17)	-0.161 (0.199)
Accountability Mechanisms	-0.036 (0.094)	-0.267 (0.15)	-0.015 (0.159)	-0.251 (0.18)
Private Complaints	-0.01 (0.052)	0.086 (0.097)	0.098 (0.094)	-0.105 (0.115)
Knowledge of Project Amount	0.697 (2.384)	-0.1 (4.318)	-1.266 (4.566)	2.334 (5.543)
Willingness to Seek Information	2.399 (2.335)	6.958 (4.453)	1.207 (4.354)	0.791 (5.265)
Quality of Accounting	0.011 (0.084)	-0.071 (0.17)	-0.075 (0.187)	0.174 (0.194)

Notes: This table replicates Table 2 for three subgroups. The column under ‘Base’ are the results as presented in Table 2. Results under ‘Projects’ are the results for villages without school rooms in 2007, “Committees” for those without any village committee or association in 2007, and ‘Inherited’ for those villages where the chief’s father’s position was inherited. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

L Registration and Mock Report

We employed a version of pre-registration of our research design. All analyses were based on hypothesis that were developed *ex ante* (in 2007) and specified without reference to evidence on treatment effects. Perhaps more critically, the core analysis was developed and coded by the research team at a time when less than 5% of data was available and without reference to actual outcomes. Instead, simulated data was analyzed and the results were written up in a “mock report”, a complete report with analysis and discussion of results, which was circulated to colleagues and posted online. This was prior to the existence of social science registries to which we could post the analysis plan. The analysis presented here differs from those described in the analysis plan in a number of ways.

First, we focus on a subset of tests, specifically we focus on the *behavioral* tests of *governance* effects. All other tests have however been implemented and are available in Humphreys et al. (2012).

Second, we made five adjustments to outcomes related to the allocation of public funds. First, financial irregularities is presented as the share of the \$1,000 instead of the dollar amount. Second, we do not report results on the number of households that claim to have received private transfers from the RAPID project because this measure is highly correlated with the standard deviation of the distributions (for the simple reason that most places had no direct beneficiaries). Third, we also chose not to condition the latter by those villages that chose distribution projects. Fourth, we added the two embezzlement measures. Fifth, related to the measure for dominance of chief’s preferences, villagers’ preferences was initially based on information from five randomly selected individuals plus from eleven additional individuals who participated in the the Step A meeting. We dropped the latter to avoid selection effects. These changes produce no substantive effect on results.

Third, we focus on participation, transparency and accountability for democratic practice. We do not focus on efficiency as this has little to do with democracy. The quality of accounting measures, however, which were part of efficiency, have been moved to the transparency section.

Fourth, for a number of outcomes we make use of mean effects analysis, as described in Section 5.3. These make for easier interpretation of otherwise multiple results.

Fifth, the analyses include block fixed effects to maximize power. We also exclude sampling weights, and focus on sample average treatment effects (SATEs). Block fixed effects are implied by the randomization strategy and focus on the SATEs reduces the risk of additional noise arising from possible errors in the measurement of village or household sizes. Results

provided in robustness tables show that these decisions are not consequential for conclusions.

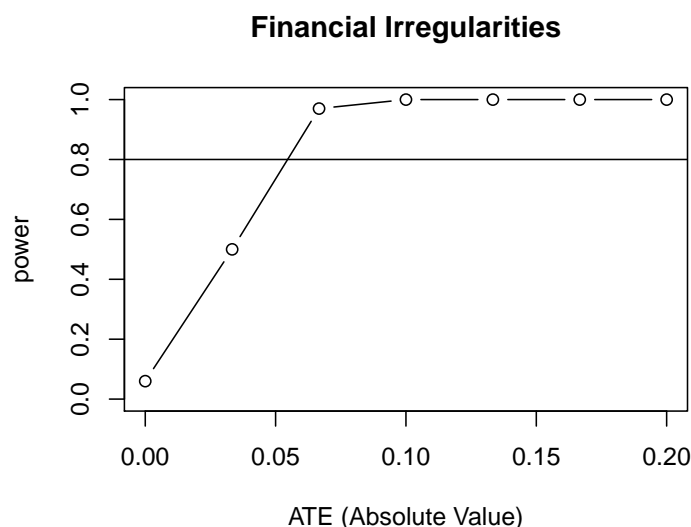
Finally, the registration document and mock report covered the main hypotheses and tests; the tests provided in Sections K and 7.2 were not elaborated in the mock report.

In all, these deviations do not substantively alter the results and all results are available as originally specified in Humphreys et al. (2012).

M Statistical Power

Figure 6 presents estimates of minimal detectable effects for our first outcome – share of funds that are not accounted for. This outcome is defined at the village level and so has, *ceteris paribus*, weaker power than measures defined at the individual level. The vertical axis shows the probability to detect an effect of a given size. The horizontal axis presents possible effect sizes. To construct this figure, we use the real outcome data variance from the control group as well as block structure and conduct simulated analysis given different possible random assignments to treatment and conjectured effect sizes. The figure thus provides the minimum treatment effect beyond which our study design would have an 80% chance to detect the treatment effect as statistically significant.

Figure 6: Minimum Detectable Effect Size for Financial Irregularities



Notes: This figure presents the power curve. On the y axis, we show the probability to detect an effect of a given size. On the x axes, we present the possible effect sizes. The figure shows the power curve for financial irregularities. To construct this figure, we use the real outcome data variance, and conduct ex-post power analysis. The figure thus provides the minimum treatment effect beyond which our study design would have an 80% chance to detect the treatment effect as statistically significant.