Institutional Change by Imitation: Introducing Western Governance Practices in Congolese Villages

Macartan Humphreys*

Raúl Sánchez de la Sierra†

Peter Van der Windt‡

November 19, 2017

Abstract

We design a system of observations of informal governance practices and elite capture in 456 villages of the Democratic Republic of the Congo, where the state has largely withdrawn. We use these measurements to estimate the effect of exposure to Western governance practices, marketed as "good governance," on the adoption of these foreign practices. To achieve this goal, we exploit a randomized intervention that provided a four year long incentivized exposure to participatory practices of governance transplanted from the West to 1,250 villages, and covered and 1.78m people. Our results suggest that such intentional, incentivized efforts to induce cultural change do not affect local decision making in the ways currently still advocated by many governments and donors. The results cast doubt on the assumptions and goals of attempts to export Western governance practice.

JEL Codes: D72; P48; D02; O17

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

---

*Columbia University, mh2245@columbia.edu.

†Corresponding author. UC Berkeley and Harvard Academy for International and Area Studies. rsanchezdelasierra@fas.harvard.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 and CARE International for their partnership in that research. Humphreys thanks the Trudeau Foundation for support while this work was undertaken. Van der Windt thanks Wageningen University.

1 Introduction

A large body of research suggests that institutions are a key driver of economic development. A core proposition is that institutions alter the incentives of political elites, producing allocations that are of greater benefit to broader sections of populations, and ultimately supportive of economic growth (Sokoloff and Engerman, 2000; Acemoglu, Johnson, and Robinson, 2001a; La Porta, Lopez-de Silanes, and Shleifer, 2008). According to North (1991), for example, "institutions form the incentive structure of a society and the political and economic institutions, in consequence, are the underlying determinant of economic performance." Although much of this research has focused on long run historical processes, international aid organizations and Western governments have taken a cue from this proposition and sought to introduce short term exposure to the practice of Western governance institutions in the hope that these will lead, through influence and adoption, to greater accountability of local elites and ultimately to the engineering of societies so that they practice governance much like the West.[1] The rapid growth of this approach reflects two weakly founded beliefs: that non-Western governance practices must be changed, and that exposure to inclusive institutional practices can lead to uptake and persistence without any changes to fundamentals.

However, despite the popularity of the underlying theory and the large amounts of development aid and government funding allocated to interventions of this form, past studies have found mixed evidence on their transformative effects and limited evidence on their mechanisms. Some studies have found evidence that, in specific areas such as behavior in public goods games, or for specific subgroups, these programs might have mild effects (Fearon, Humphreys, and Weinstein, 2009), but there is little evidence of effects on the outcomes linked to their central governance promise, such as elite decision making (Casey, Glennerster, and Miguel, 2013; Avdeenko and Gilligan, 2015; Beath, Christia, and Enikolopov, 2013; King and Samii, 2014). If indeed these interventions make little or no dint on actual behaviors, this has implications for development theory and practice. For development theory it supports the idea that institutional arrangements reflect more fundamental relations of power; for development practice it supports views that development interventions should focus on economic fundamentals instead of trying to alter political behavior hoping that these will change through imitation.

In this paper, to examine this proposition, we exploit current efforts of Western donors at achieving influence in the areas of governance through the vehicle of "Community Driven

[1] Since the 1990s, "participatory development" has become a favored model for development, and has formed a major pillar of post-conflict development. Mansuri and Rao (2013) quote a figure of $85bn in World Bank spending in the last decade alone on this broad class of interventions.

Development." In order to receive future public goods funding, villages undergo democratic training and "practice democracy" on a continuous basis during four years as they manage aid funds. To estimate the effect of training and exposure, we exploit random assignment of 1,250 villages in the Democratic Republic of Congo (DRC) to a four year long $46m Community Driven Development project, covering 1,780,000 people, and another equally sized group to the control condition of no intervention. As part of the treatment, the population is trained in the implementation of elections and accounting and accountability practices, and is subject to what amounts to social norms marketing (Paluck, 2009) and ultimately propaganda of democratic values — locally called "sensitization" campaigns.[2] The population then elects a management committee that selects a development project for the village, subject to popular approval, and receives external funds by a Western NGO to manage the development project over four years. Frequent community townhall meetings to assess performance during the four years of implementation contribute to hold the committee accountable. This provides incentives to undergo prolonged democratic process at the village level in parallel to the actual village institutions of governance. We thus take advantage of the intervention, which provides an unusually ambitious vehicle for communication and training in governance practices from societies locally perceived to be more successful, to study the plausibility of a hypothesis that enjoys strong enthusiasm but very little empirical support.

As a foundation for this paper, and to measure change of governance practices, we targeted 560 villages, half which had been previously treated by the project, with an unconditional cash grant that the village could freely decide how to spend (and also decide how to decide how to spend it). This allows us to characterize governance practices after those four years of training. Using a combination of observation of behavior in its usual environment with minimal researcher interference at different steps of the management of the unconditional cash grant, exhaustive audits by trained auditors, process data, focus groups, participation data, meetings dynamics, private project preferences, self reported kinship ties, induced dissemination of information, and privately self-reported data about knowledge of information relevant to the public, we characterize the extent to which the practices of such villages to allocate surplus are democratic. In particular, we develop measurements of participation of the typical villager, transparency of relevant information that is privately observed by the elite, accountability practices, the average level of management skills of selected representatives, social organization, and ultimately, a comprehensive set of measures aimed at understanding elite capture. This measurement strategy is a central contribution

[2]It is also made clear that such practices are the so called "good governance" practices of more successful societies.

of the paper.

Using such observations and survey data in additional villages, we characterize the governance practices of 733 villages. Across a variety of key outcomes we find no evidence that exposure to democratic practice affects informal governance practice in the average village, neither in treated villages nor in neighboring villages through spillovers. We also demonstrate that such findings cannot be explained by low statistical power, poor project 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 provide evidence that adoption of democratic practice does not follow exposure, even in such an unusually intense case of repeated exposure across years, and delivery and training of such practices by Western NGOs.

Eastern Congo is a well-suited environment to examine the adoption of democratic practice in informal governance. The state has largely withdrawn from the rural areas of the East and enjoys little legitimacy. Informal governance is often described as "captured" by traditional chiefs and exposed to corruption by state officials. This context is not unique of the Congo: a large number of post-colonial states exhibit a similar pattern (Acemoglu, Reed, and Robinson, 2014a), in which pre-colonial institutions, amplified through colonial rule, led to "decentralized despots" (Mamdani, 1996) in ways that are believed to be detrimental to development (Acemoglu, Chaves, Osafo-Kwaako, and Robinson, 2014). In addition, democratic practice, as practiced by the West, is often unknown in the rural areas.

Across a variety of outcomes we find no evidence that exposure to democratic practice affects governance practices, either in treated villages or in neighboring villages. Moreover, our results on processes have no evidence that any of the presumed mechanisms linking the intervention to governance decisions were triggered. We support this view by examining whether the failure to measure effects could be due to poor project implementation, the operation of positive spillovers, poor measurement, heterogeneous effects, elite backlash against democratization, or changes in expectations about future aid in control villages.

Our study contributes to our knowledge of institutional change, ruling out mechanisms by which institutions which are sometimes believed to be more inclusive and efficient might arise. Our choice to examine the transmission of democratic institutions relies on a unique opportunity provided by the beliefs and normative goals of Western donors which aim at promoting Western modes of governance and making poor countries seem more like the West. Since Western societies can also have institutions that are detrimental to the development of developing nations and that result from a different historical and evolutionary process, an intervention that promotes cultural influence is bound to rest of debatable ethical grounds. Despite this obvious challenge and the fact that implementers often overlook the cultural specificities of the context in which they aim to transplant institutions, the view that democratic, liberal institutions of Western nations are more inclusive (Acemoglu, Johnson, and Robinson, 2001b) and more efficient (Acemoglu and Robinson, 2012) suggests the intervention may also have some intrinsic interest.

In addition, our study makes a series of methodological contributions.

First, we develop a novel measurement 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. This improves upon standard self-reported measurements, which are subject to reporting and desirability bias (Barron, Humphreys, Paler, and Weinstein, 2009). It also improves upon laboratory games, which have been used to measure preferences and expectations, but that can lack external validity when attention to seemingly irrelevant cues can incidentally change framing and behavior (Fearon, Humphreys, and Weinstein, 2009; Lowes, Nunn, Robinson, and Weigel, Forthcoming; Haley and Fessler, 2005). On similar grounds, it improves upon the measurement of network linkages (Avdeenko and Gilligan, 2015). To observe change in governance practices, we exploit a distinct, independent unconditional cash transfer delivered in treatment and control areas after the treated areas have been exposed to four years of Western governance practices. As communities manage the unconditional cash transfer, we deploy a system of observation of community behavior that minimizes intrusion in order to reduce the risk of biasing behavior. Our strategy is similar to some of the activities studied by Casey, Glennerster, and Miguel (2013) and Beath, Christia, and Enikolopov (2013). Casey, Glennerster, and Miguel (2013) distribute a sheet of tarpaulin to villages, and enumerators later recorded how it was used (activity 3). While such activity is a significant improvement over previous literature, our measurement complements Casey, Glennerster, and Miguel (2013) in that we observe an exhaustive set of behaviors to characterize governance practice: 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 community decision, accountability - the key outcomes of democratization of governance practices. Importantly, the behaviors that we observe took place over an extended period and in a forum that was not controlled by the researchers and part of a common practice of cash transfers, outside of artificial events that might be organized by the researcher for the purpose of measurement. Tracking the use of funds allows us to measure behavior by the villagers and by the elite in a real village allocation decision, also to assess how decisions were made and ultimately who got what.

Second, our study uses to our knowledge the largest experiment of community driven development. The intervention we examine was implemented in 1,250 villages, and we employ data from 800 villages in 456 treatment clusters, which significantly reduces the likelihood of false negatives compared to previous related research. As a comparison, Fearon, Humphreys, and Weinstein (2009) examine an experiment covering 83 villages, Casey, Glennerster, and Miguel (2013) examine 236 villages, and Beath, Christia, and Enikolopov (2011) examine an intervention implemented in 217 village clusters (Casey, 2017).

Third, we employ a design-based strategy for assessing spillover effects based on randomization inference that is particularly well suited for the problem of spatial spillovers - which might arise from interventions that aim to affect expectations and which can create a strong downward bias in effect estimates. Design-based analysis of spillovers offers an approach to examine interference between units in a way that hinges on few assumptions (see Aronow and Samii (2013)).

2 The state of the literature

We examine an unusually representative (in the sense of the diversity and scope of the population covered) and ambitious version of the Community Driven approach to achieving Western influence. The project alone, with a budget of $46 million, accounted for one of the largest projects in the portfolio of the International Rescue Committee, and among the largest for the UK Department For International Development (DFID). This led the project to be called "one of the worlds largest ever randomised trials" (Hartford, 2014) and be covered in the Financial Times (Hartford, 2012). Donors were enthusiastic about the transformative effects of this project on governance culture in the villages targeted and nearby. Most decision-makers at the implementing agency and the donor agency thought it was possible or very likely that there would be "improvements" in each dimensions of governance. Half of the implementers thought it very likely that villages would manage projects in a more transparent and equitable way.[3]

Our study's goal differs from other studies examining related participatory development interventions. The related literature, reviewed in King and Samii (2014) and Casey (2017), has drawn on different versions of the participatory development package to answer different questions of interest to social science. Research in participatory development is often mistakenly associated with the "test" of a particular development model which encompasses

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

multiple components, present in varying degrees across interventions. Unlike past studies, our goal is to design a comprehensive measurement approach to characterize governance practice and ultimately transmission and persistence of governance practices. We do so taking advantage of an unusually large scale experiment that delivered exposure to democratic governance practice. While Fearon, Humphreys, and Weinstein (2009) examine whether bringing people together increases the valuation of the public good (contact hypothesis - see Tajfel (1982)), this study focuses instead on whether democratic governance practice can be adopted, leading to a change in political institutions. Unlike Beath, Christia, and Enikolopov (2013), who examine whether state aid increases state legitimacy by creating reciprocity, we explicitly focus on Western aid as the vehicle of transmission of norms and practices. This distinction is also policy relevant, since aid is often a promising lever of change when the state itself is broken.

Our study is closest to Casey, Glennerster, and Miguel (2013), but differs in important ways. First, the project we examine is of a different nature. While Casey, Glennerster, and Miguel (2013) examine a project which "reconstitutes elected district-level governments," we examine a project that introduces exposure to new democratic practice that operates outside of any pre-existing institution, and aims at inducing change in governance practice by learning and imitation rather than reinforcing the capacity of recently created institutions of governance. The intervention in Casey, Glennerster, and Miguel (2013) aims to promote the effectiveness of institutions that already exist – created recently just after the war – and that are going to continue to exist after the program, trying to foster participation and inclusion. Consistent with the objectives of the program they study, Casey, Glennerster, and Miguel (2013) examine the decision to distribute a $40 gift during one meeting. In contrast, our study is specifically designed to comprehensively measure the adoption of democratic practice, which we attempt to characterize using a wide range of measures collected around the management of a $1,000 grant which lasts 50 days. Second, while Casey, Glennerster, and Miguel (2013) and Van der Windt (2017) focus on inclusion of minorities, our study focuses instead on the effect of introducing elections and accountability in the broader governance practices of the village. In that sense, the intervention we examine, which works through electoral practice, can well induce the tyranny of the majority and increased marginalization of minority groups.

Table 1 presents the current project in light of other CDD projects, studied to examine subsets of the comprehensive promise of CDD.[4] See Casey (2017) for a general overview.

[4] We specified our core hypotheses and econometric specifications and covariates (Humphreys, Sanchez de la Sierra, and Van der Windt, 2013). See Casey, Glennerster, and Miguel (2013) for further discussion of pre-analysis plans.

Table 1: Encompassing Test of the CDD Democratization Promise

Study	Country	Research Question	Size	Effects
Fearon et al (2009)	Liberia	Bring People Together to Create Social Cohesion	83	0
Casey et al (2013)	Sierra Leone	Include Minorities in Newly Created Official Institutions	236	0
Beath et al (2011)	Afghanistan	Build State Legitimacy through Hearts and Minds	500	>0
This study	DRC	**Main:** Cultural Transmission of Democratic Practice	1,250	0
		Secondary: Create Inclusive Culture		0
		Secondary: Strengthen Social Cohesion		0
		Secondary: Legitimacy of Future Leaders		0

3 Background of governance practices in eastern Congo

In this section, we document political culture and governance practices in eastern Congo.

Following the first and second Congo wars (1996-1997 and 1998-2004), the state authority has largely retracted from most areas of eastern Congo (Raeymaekers, 2014). Instead of being exerted by the state, public authority is instead embedded in traditional chiefs (Newbury, 1991). In the areas of North and South Kivu still today, in addition, non-state armed organizations, which collect regular taxes, provide protection, and have even established their own administrations, also play a role in the governance of political institutions (Sanchez de la Sierra, 2017; Hoffmann, Vlassenroot, and Marchais, 2016; Stearns, Verweijen, and Baaz, 2013; Stearns and Vogel, 2015).

While the state is often present in the daily language of local populations, real governance is in the hands of traditional elites. So-called customary chiefs derive their power from the customs, a set of governance practices that populations have been repeating across generations, and that, by coordinating expectations, provide the basis of legitimacy (Hoffmann, 2014; Newbury, 1991). Customary chiefs are often throned following kin succession lines. When a customary chief is throned, "higher-level" chiefs, with the help of local witch doctors, invoke the tribal ancestors who support the legitimacy of the new chief. Once throned, a customary chief usually governs for life.[5] Village chiefs are below Kingdom chiefs (Chefferies). Kingdoms often follow pre-colonial ethnic boundaries.[6] Their power usually hinges on their supernatural talent, their toughness, and the perception that they are skilled leaders (Sanchez de la Sierra, 2017; Newbury, 1991). Chiefs are often evaluated on their ability to use supernatural powers to make fields grow, to secure property rights, to give protection, or to give luck, reflecting a culturally widespread belief in supernatural powers

[5]Village chiefs control a single village, while groupement chiefs control a cluster of villages. At the next level, Kingdom (Chefferie) chiefs, the "Mwamis" control an entire territory composed of tens and often hundreds of villages.

[6], although other are composed of multiple ethnic groups and their representative is appointed, rather than identified through ancestral inheritance rules

and the persistence of cultural practices in a traditional society (Nunn and Sanchez de la Sierra, 2017).

The power of chiefs relies on the ownership of land, which according to the custom, is where ancestors are buried and belongs ultimately to the chiefs. Land is then ceded to households for harvesting, often in exchange for a tax. Based on the ability to coordinate behaviors, chiefs most often use their power for the administration of justice and for taxation more generally. Chiefs often organize the provision of public goods (clearing the road, building infrastructure, and even for armed mobilization of self defense groups), drawing on an old tradition of forced labor, the Salongo.

The power of chiefs pre-dates the colonial state. However, while chiefs previously often had a role limited to coordinating expectations akin to the "leopard skin" chief in Evans-Pritchard (1969), the colonial state deeply re-inforced their power, shifting village norms towards less democratic practice. To improve their ability to govern, the Belgian administrators co-opted traditional chiefs, obtaining taxes, labor, and other mobilization through them, in exchange for the support of the coercive apparatus of the colonial state (Hoffmann, 2014). Following the colonial state, 33 years of Mobutu rule undermined the Congolese state, leading ultimately to its retraction from east Congo in practice (Schatzberg, 1997).[7] In sum, governance practices are the product of historical influence (Akyeampong, Bates, Nunn Nathan, and Robinson, 2014).

The democratic principles Western aid is keen in promoting are foreign to a diverse cultural background in rural DRC. Traditional chiefs which rule based on ethnic based custom, the product of colonial indirect rule, have emerged as the central nexus through which resources are mobilized. Given the profound impact of colonial rule in creating "decentralized despots" which then persisted (Acemoglu, Reed, and Robinson, 2014a; Mamdani, 1996), the donors that backed the intervention we study here, through the vehicle of foreign aid, explicitly aimed at exerting influence for the adoption of Western democratic practices of governance. Yet, they got confronted to a context where governance practices, culture, and the type of governance problems that might need fixing are complex and diverse, and where ethnic traditions are sticky.

[7] In 2013, Transparency International ranked the DRC the 154th most corrupt public sector out of 177 countries and territories.

4 Conceptualizing the delivery of democratic practice

In this section, we discuss how cultural transmission can change political institutions, and thus can be a vehicle for the evolution of institutions.

The sources of institutional change, and thus, how external influence can change its course, 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.[8] In the second, institutions are seen as the equilibria of a game, which endogenously constrain behavior. In the online appendix (A) 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 underlying game; the other, Strategy B, alters behavior by altering fundamentals. 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 economic growth.

Both 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 fundamentals.[9] 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 fundamentals and thus potentially amenable to equilibrium selection (Grossman and Kim, 1995; Binmore, 1998; Young, 2001; Chwe, 2000; Bidner and Francois, 2013; Acemoglu, Ticchi, and Vindigni, 2006). Different belief systems have been shown to support different equilibrium forms of social organization (Greif, 1994). Similarly, the view of constitutions as self-enforcing equilibria among administrators that coordinate to constrain the power of the ruler (González de Lara, Greif, and Jha, 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

[8]Here "rules" means the mapping from actions to payoffs, which in some applications reflects rules in a literal sense but in others reflects fundamentals more broadly.

[9]Seemingly 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).

Strategy A) emerged from outside game theory.

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, have been shown to starkly change focality of different equilibria, thus leading individuals to coordinate on new equilibria (Mehta, Starmer, and Sugden, 1992).[10]

On the other hand, research in the area of cultural evolution shows that while cultural norms often persist through generations (Alesina, Giuliano, and Nunn, 2013; Voigtlander and Voth, 2012), adoption can also occur relatively fast (Cantoni, 2011; 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? 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).[11] This is exactly what Western aid represents: the vehicle of cultural influence from more prestigious societies to less "developed" ones.

The 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, Johnson, and Robinson, 2001a; Sokoloff and Engerman, 2000).[12] Yet, it remains an empirical question whether culture, practices, and equilibrium selection, can also play an important role to explain the observed variation of institutional change.

[10] Bidner 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.

[11] Related empirical applications include "social norms marketing" in Paluck and Green (2009), and Paluck (2009)

[12] 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. Other accounts emphasize access to resources, such as subsoil resources (Ross, 2001) or aid (Nunn and Qian, 2014).

5 The Intervention: incentivized exposure to Western practice of governance

We take advantage of a large Western aid project, implemented by the International Rescue Committee during four years in 1,250 villages throughout East and South Congo with the aim of promoting the adoption of democratic practice. One of the central goals of the program was to "improve the understanding and practice of democratic governance."[13]

The goals of the program reflect a normative agenda of Western governments, funding agencies, and implementing organizations to promote Western forms of social organization, and in particular, the culture of governance, with the belief that these are better. As a reflection of this goal, the program focused on delivering "good governance," eliciting that they perceive the democratic practices they aimed to transplant to be normatively superior. Implementers emphasized that these are Western practices that countries ought to follow for development, funded by the British development agency. Protocol 76 specifies the value marketing deployed to the communities: "Improvement of good governance: practices of the transparency, accountability, representation, participatory management, inclusion of all."[14]

To achieve these goals, the donors developed one of the most representative and ambitious versions of the Community Driven Development (called "Tuungane" in DRC), due to its scale and to the stated goal of achieving cultural influence. During four years, the program disbursed $46 million of development aid to the 1,250 villages, reaching a beneficiary population of approximately 1,780,000 people, conditional on training to democratic principles and their practice in the context of the management of the project.[15]

The program followed well-defined steps. First, populations were mobilized to townhall meetings, where the objectives, the implementation agency (US NGO), and the funding government (UK) were introduced. Second, the population was trained to conduct elections, which many had never seen (14% of the chiefs in our sample are elected through elections). Third, everyone was encouraged to run.[16] Fourth, (private) elections were organized to elect a committee, whose only task would be to manage the aid fund.[17] Fifth, committee mem-

[13]In 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.

[14]Additional details on the democratic marketing messaging is in appendix 2.

[15]This amount includes the cost of some infrastructure projects across villages that are not part of this study. Figure 3 in the online appendix (C) illustrates the timing of implementation across areas.

[16]The 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.

[17]For 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

bers submitted a proposed project for popular approval in a village election.[18] Sixth, the committee was 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. Frequent elections were held where committee members could be revoked for mismanagement. Seventh, committee members received intense training to 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 project with cash or in-kind support, which aimed to create a "sunk cost" effect.

Populations exposed to Tuungane largely adopted the language of "good governance," "transparency," and "accountability" (a word whose usage was lost in French, la "redevabilité") and perceived the program as training from Western nations.

Greater detail is provided on the case in (Section K) including data suggesting that the intervention was well implemented and that, *ex ante*, practitioners expected strong results from the intervention.

6 Empirical Strategy

6.1 Random Assignment

Communities were assigned to *Tuungane* randomly, through public lotteries. All communities were first clustered and grouped geographically into 83 "lottery bins" from which clusters of villages were randomly drawn.[19] In total 600 village clusters entered the lottery, 280 were selected for treatment and the remaining 320 were in control (see Figure 1). Randomization by lottery bins achieves geographic balance across treatment and control within lottery bins. This reduces the variance of our estimates. In addition, *public* lotteries have a number of implications. First, they provide a form of informed consent on the part of communities, both those that benefit from the program and those that do not. Second, there is transparency over the selection process and this reduces concerns that one community was being unfairly favored over another. Third, 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

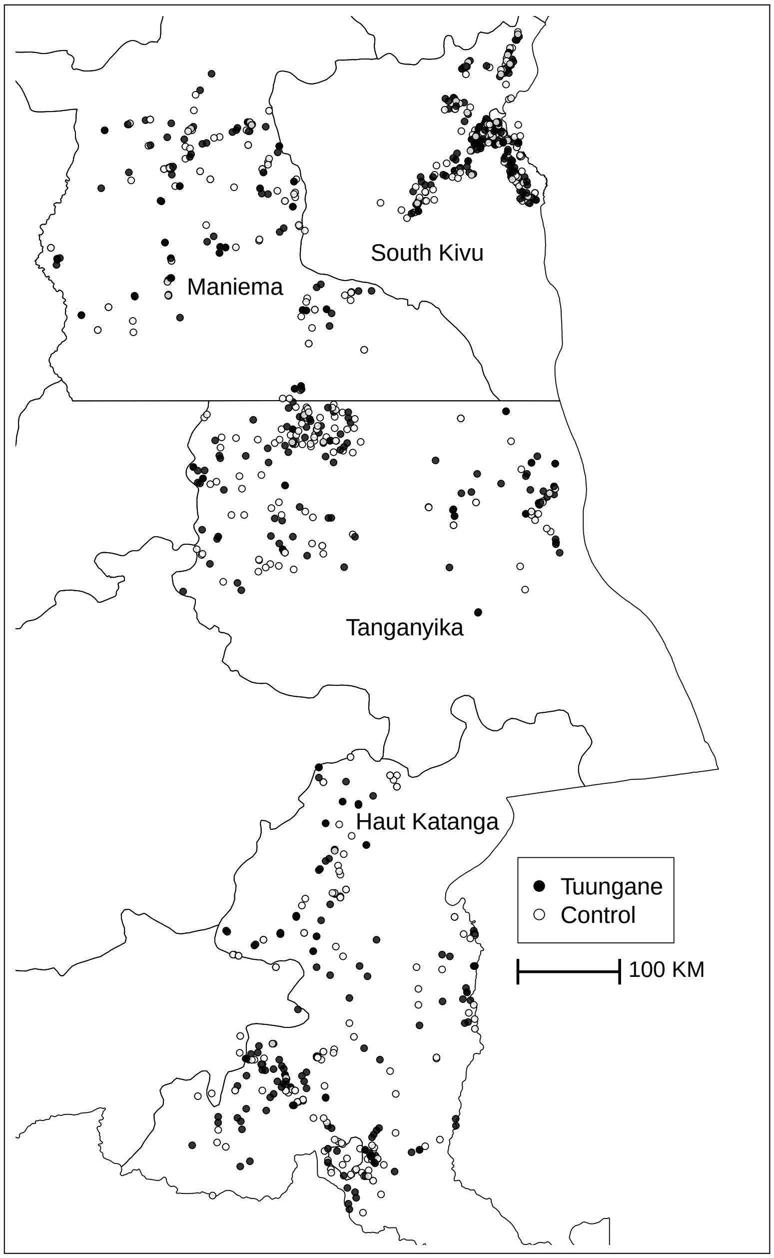
of the Election Team and TUUNGANE to ensure adequate participation. This is the only way to legitimize the elections and ensure that they are in fact a product of the community as a whole." Source: Tuungane protocols.

[18]If rejected, it was subsequently altered until a majority approved it.

[19]Lottery bins largely corresponded to chiefdoms ("Chefferies") or sectors ("Secteurs"). For simplicity, we generally use the term chiefdom for both units.

suggest that this is not likely a large concern.[20]

Figure 1: Distribution of Treatment and Control clusters



Notes Randomization was implemented at the level of blocks roughly corresponding to chiefdoms and ranging in size from 2 treatment units to 30 units. Source: Authors' drawing.

[20]We 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.

6.2 Outcomes and Measurement

The primary outcome we examine is *capture*: the extent to which the benefits provided to communities were controlled by, or benefited, the few. Measuring capture presents obvious challenges. Survey data may be biased. For example, individuals in treatment (or control) communities may try to respond in ways that please outside funders. Lab-in-the-field type measures can suffer from an interpretation challenge — with these we might observe unbiased effects but those effects may be measured on a metric with no clear interpretation.

To move beyond standard survey measures and lab-in-the-field style measures of behavior, we introduced an independent cash delivery project ("Recherche-Action sur les Projets d'Impact pour le Développement", henceforth RAPID) to assess behavioral change. As part of the RAPID process 560 villages were selected to participate in an unconditional cash transfer program in which they would receive grants of $1,000 to be used on projects that benefit the village. Of these, 280 communities had participated in the *Tuungane* program, the remaining 280 had not. Communities were unconstrained to identify and implement projects subject to minimal constraints.[21] While the RAPID project moderately encouraged distributive projects, these were not required. Importantly, the unconditional cash transfer left communities free to decide who should manage the funds and how decisions should be made. We rolled the RAPID project out in four steps (A-D) over the course of 2-3 months. The key features are described in Table 2 (the script read to the community during Step A is provided in the online appendix (See Table 6)).

To measure capture, we employ direct observations by enumerators of behavior in the village, extensive audits in each RAPID village, and large-n survey data collected during different steps of RAPID. We employ five measures of capture. One of these records the amount of funding that went missing according to our audit. Two others measure citizen perceptions of embezzlement, recorded through a direct question and a list experiment. A fourth describes the inequality in benefits received by households. The fifth captures the dominance of the chiefs' preferences in the selection of outcomes. We provide more detail on the measurement of each of these as we describe results; summary statistics for these measures can be found in supplementary materials (Section F).

In addition we collect measures on four other families of outcomes. These outcomes — *participation*, *accountability*, *efficiency*, and *transparency* — are used in Section 8 to assess whether practices encouraged by the *Tuungane* intervention, which should lead to reduced

[21] The key constraints were that some uses were ruled out if these were 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.

Table 2: 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 project funding and identify representatives (with no guidance). The population is informed that at least \$900 will be made available. Surveys are conducted with selected groups of those present during the meeting.
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 *in private* that the amount provided to villages will be \$1,000 (\$100 more than announced to the village), and of the type of audit that will be undertaken.
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 is undertaken to track the use of all funds, and measure capture, efficiency, transparency, and the accountability mechanisms 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 efficiency and equity of outcomes.

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

capture, were adopted. We describe the measures for each outcome family in the next section.

6.3 Estimation

Thanks to the experimental design, estimation is straightforward. We compare mean outcomes in *Tuungane* communities with those in control communities, which, under conditions specified by (Rubin, 1974), provides an unbiased estimate of the average treatment effect. We account for small differences in assignment propensities in different randomization bins using inverse propensity weighting and we employ sampling weights to account for differences in sampling probabilities reflecting differences in village sizes and differences in household sizes. Where individual level data is used, estimates are clustered at the level of treatment (village clusters). For transparency and consistent with our preanalysis plan, in the core analysis we include no controls. Controlling for blocks can improve efficiency Bruhn and McKenzie (2009); however in the online appendix (Section L.4) we show that effects are qualitatively unchanged when blocks are included as controls.

For some analyses we have access to multiple measures.[22] In order to generate a meaningful summary of multiple effects within each family, we follow the approach of Kling, Liebman, and Katz (2007) and create standardized indices of outcomes on related items.

7 Main Results

Our main results are given in Table 3 (see also Figure 4 in the online appendix (G) for a graphical representation of these results). On all measures, the results support the view that exposure to the grassroots democratization program had no effect on capture. The "Control" column describes the estimated level for each measure in control communities. Due to randomization, this is an unbiased estimate of the expected outcome in the absence of the program. We provide the estimated effect of *Tuungane* in the subsequent column, followed by our estimated standard error and the number of observations.

As our first measure for capture we use the share of the $1,000 grant that auditors were unable to account for. To construct this estimate, the auditors visited nearby markets to verify measures of price and quantity as listed in the accounting form. Table 3 (*"Financial Irregularities"*) presents the results: around 15% of the $1,000 could not be verified by the teams. There is no significant difference between *Tuungane* and control communities. Indeed the estimated effect is very small with a small standard error. This suggests little difference in fraudulant behavior across groups, though this does not itself mean that resources that could be accounted for were used well.

We also used two survey-based approaches to measure concerns around embezzlement of funds by village leaders. First, we asked ten individuals per RAPID village a simple direct question of whether the RAPID project went hand in hand with embezzlement — corruption, nepotism, etc. — by RAPID leaders. Table 3 (*"Embezzlement (direct)"*) shows that around 13% of respondents in control communities report this to be the case. Results are similar in *Tuungane* communities. To overcome possible measurement error resulting from directly asking these questions, we also make use of a list experiment. 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.[23] Because of aggregation, the respondent can be assured that nobody

[22]This raises concerns about interpretation. For example, it may be that all measures trend positive, but none is individually statistically significant. In such a case it is possible that effects are jointly significant across the family of measures. Conversely, it may be that by chance a certain measure is significant in a family while most are not, or even trend in the wrong direction. In such cases it is possible that there are no significant effects across the family of measures.

[23]The sensitive item was: "There was embezzlement in the village by the RAPID leaders (corruption,

knows the answer to the sensitive question. We take the difference between the average response in both groups as our measure for embezzlement.[24] Table 3 (*"Embezzlement (list experiment)"*) presents the results. In control areas, around 14% of respondents indicate that embezzlement took place. We find no statistically significant difference in *Tuungane* areas, although for this measure, which draws on an interaction term, the estimated effect points in the wrong direction and is not measured with great precision.

Table 3: Main Results

	Control	Effect	(se)	N
Financial Irregularities	0.148	-0.004	(0.021)	394
Embezzlement (direct)	0.129	0.006	(0.024)	1807
Embezzlement (list experiment)	0.135	0.107	(0.133)	1608
Inequality of (Private) Benefits	2.566	0.391	(0.593)	409
Dominance of Chief's Preferences	0.082	0.013	(0.043)	2111

Notes: For a more complete discussion on each measure, see Humphreys, Sanchez de la Sierra, and Van der Windt (2012). For "Embezzlement (list experiment)" and "Dominance of Chief's Preferences" we estimate $Y = \beta_0 + \beta_1 X + \beta_2 T + \beta_3 X \times T$, and report β_1 in the Control column and β_3 in the Effect column, where X is the sensitive item and the chief, respectively. All analyses employ propensity score weights and clustering of standard errors at the level of randomization clusters. Reported number of observations refer to survey respondents (rows 2 and 3), villages (1 and 4), and the chief plus randomly sampled village members (5 per village) (row 5). $*p \leq 0.10, **p \leq 0.05, ***p \leq 0.01$.

Our fourth measure relates to the distribution of economic benefits from the RAPID project. We asked ten individuals per RAPID village a simple direct question of whether their household received something directly from the RAPID project. Subsequently, the value was calculated together with our enumerator. As a direct measure of capture of economic benefits, we compute the standard deviation of the distributions that took place

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 project; c) The population would have liked RAPID to support the project with technical knowledge.

[24]In the analysis in Table 3 this is calculated by taking the coefficient on the interaction between treatment and an indicator for a long list; in later village level analyses the dependent variable is defined as the within village difference between long list and short list responses.

(in dollars) to represent the average difference in the amount received between two randomly selected villagers. On average, in control communities this standard deviation is around \$2.6, and indistinguishable from *Tuungane* communities. We note that in villages in which public goods were produced and no distributions with cash value were made this standard deviation is zero.

Finally, we provide results from a behavioral measure that captures the extent to which actual decisions disproportionately reflect the preferences of the village chiefs and not those of the villagers. Our measure compares the predictive power of the chief's preferences to those of a random sample of five villagers. We find in Table 3 (*"Dominance of chief's Preferences"*) that in all areas the project realization (obtained during Step D) coincides better with the stated preferences (taken during Step A before the village meeting) of the chief than those of the villagers. In control areas the chief's prior preferences are eight percentage points more likely than those of a randomly selected villager to coincide with actual projects. Chiefly dominance is around one percentage point higher in *Tuungane* areas (around 9%). We thus find evidence of chief dominance in all areas but little evidence that the *Tuungane* program reduces the strength of this dominance.

In conclusion, we find no evidence that exposure to democratic practice has an impact on governance outcomes. Across a set of measures we find that our estimated effects go in the wrong direction; and many of our nulls are estimated with reasonable precision.

8 Why No Effect?

We examine two sets of reasons for the lack of an observable effect. In Section 8.1 we examine mechanisms. In particular, we seek to understand whether the propensity to capture benefits was unaffected because prior outcomes were unaffected, or perhaps despite the fact that prior outcomes were affected. In section 8.2, we examine whether alternative mechanisms played a role at biasing the effect downwards to zero.

8.1 Mechanisms

Most pertinent to both the theoretical accounts and the policy motivations, we focus on four channels through which demonstration effects may operate. The first is that exposure affects propensities for individuals to participate in politics. Greater participation provides a channel through which decisions favor broader groups. Two other channels relate to behaviors and practices around transparency and accountability. If accountability and transparency practices from the *Tuungane* intervention are inherited by communities when they face the RAPID decision making problem they may facilitate citizen control of elites. Finally, we examine capacity rather than politics. Corruption and incompetence are sometimes observational equivalents and increased benefits can accrue to broader populations simply by virtue of the fact that leaders are better able to manage funds or assess how they should be allocated. While these four channels are not exhaustive (notably we do not focus here on channels that operate through changing values), they do correspond to major features identified by literatures on allocative decision making and are ones that were identified in advance by policy experts to justify interventions of this form.

We examine these four next. We highlight that our examination of channels focuses for the most part (with the exception of measures of willingness to seek information and the speed of information transmission) on behavior observed during the RAPID process. Thus we assess whether the institutional intervention altered participation, say, which in turn may have affected capture. It is possible, however, that the intervention could have affected willingness to participate to prevent capture but that this does not translate into greater participation because of the effectiveness of the threat. Even still, this approach does allow us to assess whether the practices and behaviors introduced by the project with the goal of altering outcomes were adopted in subsequent political decision making.

8.1.1 Participation

A set of participation measures captures the extent to which communities adopted practices intended to provide formal scope for broad input to decision making during RAPID. We begin with an examination of villagers' attendance and patterns of social interaction in the first village meeting (Step A). Table 4 ("*Meeting Attendance*") shows that in control communities on average 130 adults participated in the public meeting, two more than in *Tuungane* communities, a very small difference which is not statistically significant. These meetings were attended by one of our enumerators who recorded patterns of social interaction. The second row ("*Interventions in Meetings*") 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 73% of the interventions. The third row ("*Dominance of Men in Discussion*") shows that the patterns of male dominance of social interactions are indistinguishable across treatment and control.[25]

Participation in village meetings may be easy to manipulate at no cost, however. We

[25]We find similar results for dominance of the chief and elderly.

therefore provide an additional measure of participation that focuses on the process in which the committees were selected. Between Step A and B, RAPID communities were required to select both a committee and a project as part of the terms of receiving funds. After leading two simultaneous focus groups, one with members of the committee and a second with ordinary villagers during Step B of the RAPID process, our enumerator teams coded the selection process as either electoral, through lottery, by consensus, imposed by the chief or elders, other or unknown. Approximately 43% of committees and 31% of projects were coded as selected through election, and 71% of committees and 73% of the projects were selected through either election, lottery or consensus. We combine these four indicators as a (normalized) composite measure, which by construction equals zero in control areas and has a standard deviation equal to one. Table 4 ("*Participatory Selection Methods*") shows that we find no evidence that participation in *Tuungane* leads to greater adoption of participatory processes in the selection of the committee or projects. The estimated effect is below one tenth of a standard deviation.

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 ("*Committee Composition*"). There was no constraint placed on the composition of these committees other than size (at least 2 members and no more than 8). Our composite measure includes the number of women, the number of men, the total size, and the share of women on the committee. We find a strong tendency towards male domination of committees: of 452 committees, 28 had gender parity, two had more women than men, and the rest had more men than women. On average about 1 committee member in 7 was a woman (18% in control; 20% in treatment). Again, on the composite index we find no statistically significant difference between *Tuungane* treatment and control communities.[26]

8.1.2 Accountability

A set of accountability measures are used to assess whether exposure to good governance practices led to adoption of similar mechanisms during RAPID.

First, we combined survey and focus group data to construct a measure of what kinds of accountability mechanisms were organized to oversee committees. We find that 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

[26] Looking at the number of women and the share of women individually, we do find evidence that the *Tuungane* program had an impact though significance is lost when we examine the index.

Table 4: Mechanisms

	Control	Effect	(se)	N
Participation				
Meeting Attendance	130.482	-1.983	(7.367)	455
Interventions in Meeting	14.695	-0.391	(0.509)	457
Dominance of Men in Discussion	72.795	0.651	(1.385)	457
Participatory Selection Methods	0	0.07	(0.093)	451
Committee Composition	0	0.076	(0.091)	452
Accountability				
Accountability Mechanisms	0	0.056	(0.107)	413
Private Complaints	0	0.022	(0.072)	3652
Transparency				
Knowledge of Project Amount	38.639	1.462	(3.207)	3694
Willingness to Seek Information	37.697	3.839	(3.277)	1406
Efficiency				
Quality of Accounting	0	0.013	(0.105)	399
Information Transmission	9.644	-1.339	(1.565)	3800

Notes: For a more complete discussion on each measure, see Humphreys, Sanchez de la Sierra, and Van der Windt (2012). Outcome measures in rows 4 to 7 and row 10 are indices and by construction equal zero in control areas. For "Information transmission" we estimate $Y = \beta_0 + \beta_1\text{RAPID} + \beta_2 T + \beta_3\text{RAPID} \times T$, and report β_1 in the Control column and β_3 in the Effect column. All analyses employ propensity score weights and clustering of standard errors at the level of randomization clusters $*p \leq 0.10, **p \leq 0.05, ***p \leq 0.01$.

(such as a distinct committee) had been put into place, and another 13% indicated that the committee had been required to report its actions to the community as a whole. As the composite measure in Table 4 indicates ("*Accountability Mechanisms*"), *Tuungane* did not lead to a greater propensity to put accountability mechanisms into place.[27]

Second, during during Step D, we asked ten randomly selected respondents to indicate whether or not they agreed with thirteen pre-selected complaints. As calculated by an index of the average propensity of villagers to issue complaints, results in Table 4 ("*Private Complaints*") suggest that levels of complaint are no higher in *Tuungane* areas than in control. Note that it could be that they are no higher because project management was better, though as we saw above there is little evidence to support this.[28]

8.1.3 Transparency

Two further measures of information transmission are used to assess whether the principles of transparency emphasized by the *Tuungane* intervention were subsequently adopted.

To measure information delivery by the elite to the villagers, the enumerator teams informed the community in the first meeting (Step A) that $900 *or more* will be made available. In all communities, however, the self-identified committee received $1,000 one week later during Step C. We are able to measure information delivery by interviewing a random sample of villagers about the amount of the RAPID grant to see whether they report the dollar amount told to them by us in Step A, or rather the true amount, as known by committee members. Table 4 ("*Knowledge of Project Amount*") shows that, on average, 39% of all respondents (and 56% of those respondents that gave an answer) report the correct answer of $1,000. However, we find no evidence that there is a difference between treatment and control communities.

To assess the willingness and ability of randomly selected villagers to obtain relevant information about the management of public resources, we went a step further asking a sample of 1,406 respondents to seek out information on fund use in their communities. From Table 4 ("*Willingness to Seek Information*") we see that approximately 38% 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

[27] The composite measure includes nine variables: three measures (external accountability measure, committee, or any mechanisms) from three different sources (focus group with the RAPID committee, interview with two RAPID committee members, interview with 10 random villagers). Note the drop in observations when moving from our measures for participation to those for accountability and efficiency. The reason is the expulsion of our enumerator teams from Maniema. This happened after most of the Step A data was collected, but before much of the Step D data was collected. Our measures for participation (accountability and efficiency) build on Step A (D) data. In the online appendix (L.1) we discuss attrition in detail.

[28] We 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 endogeneous.

(75), that the exercise is strange to them (50), that the husband of the respondent refuses or would refuse the collection of this information (13), and other reasons (192). This suggests that accessing basic financial information is challenging. There is no significant difference between treatment and control communities.

8.1.4 Efficiency

Finally, we collect a set of efficiency measures to capture changes in capacity and effort by elites and the broader population, as well as the rate of information transmission in villages.

RAPID committees were expected to keep an accounting form, and record the total amount made available for the project (out of $1,000) and the expenditures made. We use the presence of this form, and information on whether it was completed as an indicator of efficient implementation. We find that on average, in 82% of the villages, committees had their accounting form present upon arrival of the audit team during Step D. Approximately 78% of the funds were formally accounted for as calculated by the RAPID Committee (and 83% when calculated by the audit teams). In addition, 56% of the money the committee made available for the RAPID project (of the $1,000) was justified by receipts, and 46% was justified with receipts deemed credible by the auditing team. Table 4 ("*Quality of Accounting*") 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.

Finally, we generate a behavioral measure of the extent to which the community can function efficiently outside of the RAPID process by examining information transmission among villagers. For this we delivered public health information on hygiene and diarrhea to a random sample of villagers in the first visit. We then returned to these villages (Step D) and interviewed the same villagers, as well as a random sample of villagers who were not yet visited. We compare the answers provided by these additional villagers in a health information test to the answers provided by villagers in another 396 randomly selected villages in which no information was delivered but the tests were implemented simultaneously. This allows us to assess the rate of information spillovers within villages, and compare it between treatment and control. As expected, we find that those living in villages where we distributed the information to other people score ten points higher on a set of questions related to the public health information we provided. This result indicates that information spreads. However, as Table 4 ("*Information Transmission*") shows, we see no evidence that *Tuungane* had an impact on this information transmission. The RAPID effect is smaller by 1.3 points in *Tuungane* villages than in non-*Tuungane* villages, suggesting that *Tuungane* villages may do marginally worse at information transmission.

In summary, we find no evidence that any of the suggested intermediate mechanisms — *participation*, *accountability*, *efficiency*, and *transparency* — was triggered by the intervention. This can be one reason why we find that exposure to a grassroots democratization program had no effect on capture.

8.2 Alternative explanations

Another possibility is that our results and measures suffer from bias of different forms. Two especially pernicious types of bias are social desirability bias and spillover bias.

Here we assess whether spillovers to neighboring areas are present, as these could lead to a downward bias in estimated treatment effects. We then assess whether weak results could be due to differential desirability bias. In addition, in the online appendix we consider the possibility that results reflect a bias due to a short term elite response to the intervention (online appendix J), and then explore concerns related to data missingness, compliance, treatment heterogeneity and specification biases (online appendix L). We also explore, in appendices, survey based information that addresses the concern that the case itself was a poor case for finding evidence of effects or that the intervention was poorly implemented (Section K).

8.2.1 Spillovers

If *Tuungane* produced positive effects beyond treatment communities, the null results could be the result of positive spillovers. We use a design based strategy to assess the presence of spillovers.[29] That is, 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.[30] 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 to recover these propensities, since they are determined by our original randomization. To calculate these propensities we randomly re-allocate the *Tuungane* treatment 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,

[29] See Gerber and Green (2012), Chapter 8, for more details. Another, more basic approach that uses the distance to the nearest *Tuungane* village as an alternative treatment (conditional on lottery bin and shortest distance to any village) in order to capture spillover effects, produces similar results.

[30] Note that for the spillover analysis missing data affects both the set of units in the study *but also the measures of exposure to spillovers*. Our results assess the effects of being close to a treatment village for which we have location data. We ignore this distinction 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).

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 any value of x. We then generate estimates of treatment effects by comparing outcomes in each combination of conditions with inverse propensity score 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).[31]

We conduct our analysis for both a 5km radius spillover treatment and a 20km radius spillover treatment. We highlight (and illustrate in Figure 6 in online appendix H) that when we examine different conceptualizations of the treatment effect we *simultaneously alter our samples*. The intuition is that 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 109 (199) units, with only 20 villages being in both groups.[32] Finally, we note that 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 does not (Bowers and Fredrickson, 2013).

The results (presented in Table 10 in online appendix L) shows no evidence that exposure to *Tuungane* had spillover effects on good governance (or direct effects, once we take account of possible spillovers). This conclusion is also supported by the following considerations. First, even if there were spillovers we would expect them only to attenuate the estimated effect, not to eliminate or produce effects of the wrong sign.Second, randomization was implemented at the level of clusters of VDCs, hence most treated villages are surrounded by treated villages and most control villages by control villages, which limits the scope for

[31] That 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 under the sharp null. We compare the actual estimated effect to this distribution and estimate how likely it is we would have obtained results as strong or stronger than our estimated effects under the sharp null. See also: Barrios, Diamond, Imbens, and Kolesár (2012).

[32] 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 109 villages: 19 neither direct nor indirect, 35 not direct but indirect, 30 direct and not indirect, and 25 direct and indirect. At a 5km radius these categories total, respectively, 44, 55, 47 and 53 (summing up to 199).

spillovers to control areas. Third, populations in control areas report low levels of knowledge about *Tuungane*.

8.2.2 Differential Desirability Bias

Another possible bias we consider is that control communities, expecting future conditional aid disbursement, and having not received the infrastructure funds, may have managed the cash in a more democratic manner in order to please future donors. This could be the case if the RAPID cash delivery project was perceived as linked to the donor community, despite our best efforts to uncouple them.[33] We use a small experiment embedded in our endline survey to shed light on this possibility.

To assess directly whether villagers strategically displayed behaviors aligned with their beliefs about expectations of development donors, we introduced a survey variation in which we asked a randomly selected set of 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 one randomly selected subgroup the question was preceded by the statement "Many NGOs in the region think that election are not the best way to choose community representatives when it comes to an appointment with technical responsibilities"; another subgroup 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.[34]

We find strong evidence for a social desirability bias: in both groups individuals are 19 - 23 percentage points more likely to answer 'yes' following a positive prompt (standard error: ca 0.025 / 0.027) (full results are provided in Table 9 in online appendix I). However, we do not find evidence that this bias is affected by exposure to *Tuungane*. The difference between the two groups is small but is stronger (though not significantly so) for the *Tuungane* group.

[33] More 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 project RAPID 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 disbursing funds. Moreover, the IRC and CARE International employees that visited villages to distribute the 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.

[34] More 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.

Therefore, social desirability bias among control communities is unlikely to drive our result.

9 Conclusion

Participatory development is still often emphatically used as a strategy to promote persistent change in governance practices across the world, so that they resemble to those of the West. This enthusiasm often ignores the possibility that existing governance practices may have a functional purpose, as well as their stickiness due to the persistence of culture and of power relations. We examined the effects of an unusually ambitious and representative participatory development project in the Democratic Republic of the Congo on subsequent democratic practice. Our findings suggest that four years of exposure to new practices that are presented to be "better" and like in the West does not lead to subsequent change.

These findings contribute to the literature on the political economy of development and to development practice. First, our findings suggest that behavioral change of local elites and of the general population stemming from demonstration effects and foreign exposure to new practices is not likely even after four years of exposure. This resonates with the approach of scholarship studying long-run development and institutional change (Akyeampong, Bates, Nunn Nathan, and Robinson, 2014).

Our results do not suggest that local elections or electoral processes do not matter. A series of results suggest that they do, including observational and experimental results. Martinez-Bravo, Padro i Miguel, Qian, and Yao (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 it can worsen local collective action. These studies look at settings that differ from ours in a set of related ways. First, these studies focus primarily on the effects of institutions *on the decisions made under these institutions*. Second, they focus on cases in which persistence is exogenous — guaranteed by an external group. And, third, by looking at institutions that connect communities to larger structures such as national governmental structures — these studies examine settings where institutions provide a clear mechanism through which change in governance is achieved.

Second, our findings suggest the need to rethink the strategies employed by governments in developing countries and donors to engage with institutions with low capacity, corruption, or weak accountability. Current donor-driven approaches to render decision-making more inclusive by short/medium-term interventions, which do not change the economic fundamentals, may serve great the population as they last, but their transformative promise has no empirical support. Of course, we posit, when sectors of the population are likely to have vested interests, institutional change may itself require a change in the allocation of economic power.

The logic of exogenous institutional change is that external action helps shift populations from one equilibrium to another believed to be better. This presupposes that these populations are in a bad 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 frequently adopt a chiefs-as-despots model and view rural institutions as captured by traditional elites. Moreover CDD programs are often motivated by a claim that past conflict breaks local accountability by undermining the ability of communities to cooperate. Our data is inconclusive about this assumption but suggests it might be exaggerated. On the one hand, we uncover financial irregularities affecting 15% of funds across areas, a domination of chiefly preferences across areas, and less than 50% knowledge of project amounts. On the other hand, we found that very few respondents felt they had a right to seek out information on fund usage in their communities. However, we also found higher baseline levels of general participation, public information, and perceived legitimacy of existing decision making mechanisms across these communities than often otherwise expected.

We cannot conclude from this study that the institutional claim is incorrect. Indeed, when phrased as a possibility claim, it resists falsification. The implications of our findings relate to the application of the claim rather than the claim itself: the institutional logic provides grounds for optimism that short term interventions may produce changes in expectations that produce large changes in behavior without any accompanying changes in fundamentals. Our results, and others, speak against that inference in cases where development actors have placed great confidence in it.

References

Acemoglu, D., I. N. Chaves, P. Osafo-Kwaako, and J. A. Robinson (2014): “Indirect Rule and State Weakness in Africa: Sierra Leone in Comparative Perspective,” *NBER Working paper*, (20092). 3

Acemoglu, D., S. Johnson, and J. Robinson (2001a): “The Colonial Origins of Comparative Development: An Empirical Investigation,” *American Economic Review*, 91(5), 1369–1401. 1, 10

Acemoglu, D., S. Johnson, and J. A. Robinson (2001b): “The Colonial Origins of Comparative Development : An Empirical Investigation,” *The American Economic Review*, 91(5), 1369–1401. 4

Acemoglu, D., T. Reed, and J. a. Robinson (2014a): “Chiefs: Economic Development and Elite Control of Civil Society in Sierra Leone,” *Journal of Political Economy*, 122(2), 319–368. 3, 8

Acemoglu, D., T. Reed, and J. A. Robinson (2014b): “Chiefs: Elite Control of Civil Society and Economic Development in Sierra Leone,” *Journal of Political Economy*, 122(2), 319–368. 55

Acemoglu, D., and J. A. Robinson (2012): *Why Nations Fail: The Origins of Power, Prosperity and Poverty*. Crown, New York, 1st edn. 4

Acemoglu, D., D. Ticchi, and A. Vindigni (2006): “Emergence and Persistence of Inefficient States,” Working Paper 12748, National Bureau of Economic Research. 9

Akyeampong, E., R. Bates, Nunn Nathan, and J. Robinson (2014): *Africa’s Development in Historical Perspective*. Cambridge University Press, Cambridge. 8, 27

Alesina, A., P. Giuliano, and N. Nunn (2013): “On the Origins of Gender Roles: Women and the Plough,” *Quarterly Journal of Economics*, 128(2), 469–530. 10

Aoki, K., and M. W. Feldman (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. 10

Aronow, P. M., and C. D. Samii (2013): “Estimating Average Causal Effects Under General Interference,” *Working paper*. 5

Avdeenko, A., and M. J. Gilligan (2015): "International Interventions to Build Social Capital," *American Political Science Review*, 109(3), 427–449. 1, 4

Baldwin, K., and E. Mvukiyehe (2015): "Elections and Collective Action: Evidence from Changes in Traditional Institutions in Liberia," *World Politics*, 67(04), 690–725. 27

Barrios, T., R. Diamond, G. W. Imbens, and M. Kolesár (2012): "Clustering, Spatial Correlations and Randomization Inference," *Journal of the American Statistical Association*, 107(498), 578–591. 25

Barron, P., M. Humphreys, L. Paler, and J. M. Weinstein (2009): "Community Based Reintegration in Aceh," *Indonesian Social Development Papers*, 12. 4

Beath, A., F. Christia, and R. Enikolopov (2011): "Elite Capture of Local Institutions: Evidence from a Field Experiment in Afghanistan," *Working paper.* 5

——— (2013): "Do Elected Councils Improve Governance? Experimental Evidence on Local Institutions in Afghanistan," *World Bank Policy Research Working Paper Series*, (6510). 1, 4, 6

Bidner, C., and P. Francois (2013): "The Emergence of Political Accountability," *Quarterly Journal of Economics*, 128(3), 1397–1448. 9, 10

Binmore, K. G. (1998): *Game Theory and the Social Contract: Just Playing.* MIT Press, Boston. 9

Bowers, J., and M. M. Fredrickson (2013): "Reasoning about Interference Between Units," *Political Analysis*, 21(1), 97–124. 25

Boyd, R., and P. Richerson (2002): "Group Beneficial Norms Can Spread Rapidly in a Structured Population," *J Theor Biol*, 215, 287–296. 9, 10

Bruhn, M., and D. McKenzie (2009): "In Pursuit of Balance: Randomization in Practice in Development Field Experiments," *American Economic Journal: Applied Economics*, 1(4), 200–232. 15

Cantoni, D. (2011): "Adopting a new religion: The case of Protestantism in 16th Century Germany," Economics Working Papers 1265, Department of Economics and Business, Universitat Pompeu Fabra. 10

Casey, C. (2017): "Radical Decentralization: Does community driven development work?," *Working paper.* 5, 6

Casey, K., R. Glennerster, and E. Miguel (2013): "Reshaping Institutions: Evidence on Aid Impacts using a Preanalysis Plan," *Quarterly Journal of Economics*, 127(4), 1755–1812. 1, 4, 5, 6

Chattopadhyay, R., and E. Duflo (2004): "Women as Policy Makers: Evidence from a Randomized Policy Experiment in India," *Econometrica*, 72(5), 1409–1443. 9

Chwe, M. S.-Y. (2000): "Communication and Coordination in Social Networks," *Review of Economic Studies*, 67(1), 1–16. 9

Evans-Pritchard, E. E. (1969): *The Nuer*. Oxford University Press, London. 8

Fearon, J. D., M. Humphreys, and J. M. Weinstein (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. 1, 4, 5, 6

Fisher, R. A. (1935): *The Design of Experiments*. Oliver and Boyd, London. 25

Fujiwara, T. (2015): "Voting technology, political responsiveness, and infant health: evidence from Brazil," *Econometrica*, 83(2), 423–464. 9

Gerber, A. S., and D. P. Green (2012): *Field Experiments: Design, Analysis, and Interpretation*. W.W. Norton, New York City. 24

Giuliano, P., and N. Nunn (2017): "Understanding Cultural Persistence and Change," . 10

González de Lara, Y., A. Greif, and S. Jha (2008): "The Administrative Foundations of Self-Enforcing Constitutions," *American Economic Review*, 98(2), 105–109. 9

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. 9

Greif, A., and D. D. Laitin (2004): "A Theory of Endogenous Institutional Change," *American Political Science Review*, 98(4), 633–652. 9, 36

Grossman, G. (2014): "Do Selection Rules Affect Leader Responsiveness? Evidence from Rural Uganda," *Quarterly Journal of Political Science*, 9(1), 1–44. 27

Grossman, H. I., and M. Kim (1995): "Swords or Plowshares? A Theory of the Security of Claims to Property," *Journal of Political Economy*, 103(6), 1275–1288. 9

Haley, K. J., and D. M. Fessler (2005): "Nobody's Watching? Subtle Cues Affect Generosity in an Anonymous Economic Game," *Evolution and Human Behavior*, 26(3), 245–256. 4

Hamilton, W. D. (1964): "The Genetical Evolution of Social Behavior," *Journal of Theoretical Biology*, 7(1), 1–16. 50

Hartford, T. (2012): "Cash Delivery on the World's Poorest, https://www.ft.com/content/7ce1b356-f6f9-11e1-9dff-00144feabdc0," *Financial Times*. 5

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/," Discussion paper. 5

Herbst, J. (2014): *States and Power in Africa*. Princeton University Press, Princeton. 10

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. 7, 8

Hoffmann, K., K. Vlassenroot, and G. Marchais (2016): "Taxation, Stateness and Armed Groups: Public Authority and Resource Extraction in Eastern Congo," *Development and Change*, 47(6), 1434–1456. 7

Humphreys, M., R. Sanchez de la Sierra, and P. Van der Windt (2012): "Social and Economic Impacts of Tuungane: Final Report on the Effects of a Community Driven Reconstruction Program in Eastern Democratic Republic of Congo," . 17, 21, 58, 59

——— (2013): "Fishing, Commitment, and Communication: A Proposal for Comprehensive Nonbinding Research Registration," *Political Analysis*, 21(1), 1–20. 6

King, E., and C. D. Samii (2014): "Fast-Track Institution Building in Conflict-Affected Countries? Insights from Recent Field Experiments," *World Development*, 64(1), 740–760. 1, 5

Kling, J. R., J. B. Liebman, and L. F. Katz (2007): "Experimental Analysis of Neighborhood Effects," *Econometrica*, 75(1), 83–119. 16

La Porta, R., F. Lopez-de Silanes, and A. Shleifer (2008): "The Economic Consequences of Legal Origins," *Journal of Economic Perspectives*, 46(2), 233–285. 1

Lowes, S., N. Nunn, J. A. Robinson, and J. Weigel (Forthcoming): "The Evolution of Culture and Institutions: Evidence from the Kuba Kingdom," *Econometrica*. 4

Mamdani, M. (1996): *Citizen and Subject: Contemporary Africa and the Legacy of Late Colonialism*. Princeton University Press, Princeton. 3, 8

Mansuri, G., and V. Rao (2013): *Localizing Development: Does Participation Work?* World Bank Policy Report. 1

Martinez-Bravo, M., G. Padro i Miguel, N. Qian, and Y. Yao (2011): "Do Local Elections in Non-Democracies Increase Accountability? Evidence from Rural China," *NBER Working paper*, 16948. 27

Mead, M. (1968): *New Lives for Old: Cultural Transformation -Manus, 1928-1953*. 10

Mehta, J., C. Starmer, and R. Sugden (1992): "An Experimental Investigation of Focal Points in Coordination and Bargaining: Some Preliminary Results," in *Decision Making under Risk and Uncertainty*, pp. 211–219. Springer. 10

Newbury, D. (1991): *Kings and clans: Ijwi Island and the Lake Kivu Rift, 1780-1840*. University of Wisconsin Press. 7

North, D. C. (1991): "Economic Performance Through Time," *The American Economic Review*, 84(3), 359–368. 1

Nunn, N., and N. Qian (2014): "US Food Aid and Civil Conflict," *American Economic Review*, 104(6), 1630–1666. 10

Nunn, N., and R. Sanchez de la Sierra (2017): "Why Being Wrong Can Be Right: Magical Warfare Technologies and the Persistence of False Beliefs," *American Economic Review, Papers and Proceedings*, 107(5), 582–87. 8

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. 2, 10

Paluck, E. L., and D. P. Green (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. 10

Raeymaekers, T. (2014): *Violent Capitalism and Hybrid Identity in the Eastern Congo: Power to the Margins*. 7

Ross, M. L. (2001): "Does Oil Hinder Democracy?," *World Politics*, 53(3), 325–361. 10

Rubin, D. B. (1974): "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies," *Journal of Educational Psychology*, 66(5), 688–701. 15

Sanchez de la Sierra, R. (2017): "On the origins of the state: Stationary Bandits and Taxation in Eastern Congo," *Unpublished*. 7

Schatzberg, M. G. (1997): "Beyond Mobutu: Kabila and the Congo," *Journal of Democracy*, 8(4), 70–84. 8

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. 9

Shepsle, K. A. (2006): "Rational Choice Institutionalism," *The Oxford Handbook of Political Institutions*, pp. 23–38. 9

Sokoloff, K. L., and S. L. Engerman (2000): "History Lessons Institutions, Factor Endowments, and Paths of Development in the New World," *Journal of Economic Perspectives*, 14(3), 217–232. 1, 10

Stearns, J., J. Verweijen, and M. Baaz (2013): *The national army and armed groups in the eastern Congo: untangling the Gordian knot of insecurity*, Usalama Project. Rift Valley Institute. 7

Stearns, J., and C. Vogel (2015): *The Landscape of Armed Groups in Eastern Congo*. Center on International Cooperation at New York University, New York. 7

Tajfel, H. (1982): "Social Psychology of Intergroup Relations," *Annual Review of Psychology*, 33, 1–39. 6

Van der Windt, P. (2017): "Can Development Aid Empower Women?," *Working Paper*. 6

Voigtlander, N., and H.-J. Voth (2012): "Persecution Perpetuated: The Medieval Origins of Anti-Semitic Violence in Nazi Germany*," *The Quarterly Journal of Economics*, 127(3), 1339. 10

Young, H. P. (2001): *Individual Strategy and Social Structure: An Evolutionary Theory of Institutions*. Princeton University Press, Princeton. 9, 10

Online Appendix

A Theoretical framework

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 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 game unchanged.

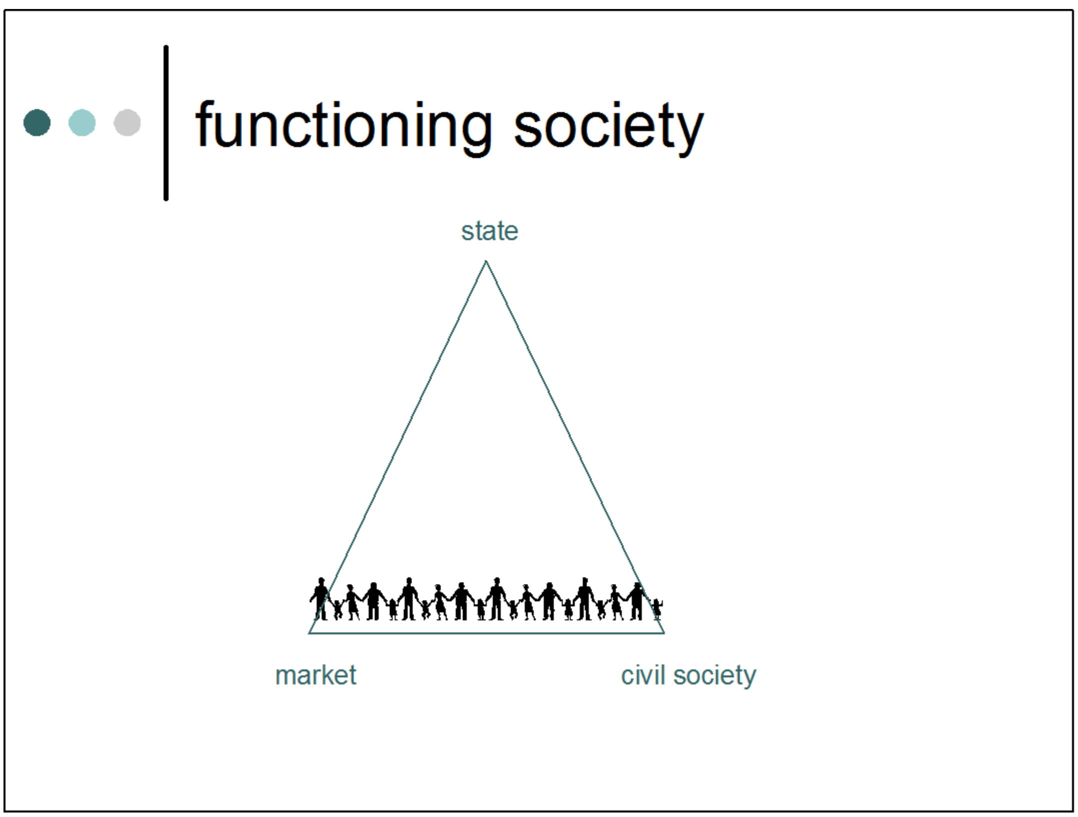
Strategy B. Consider now a second intervention in which the third party guarantees W a return of $d_W = 0.75$ instead of $d_W = 0.5$ in the event of cooperation failure. This is a structural change and has a real effect on W's bargaining position. It means that W can now do better playing D in all periods 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 Democratic Marketing

The following image, taken from Tuungane protocols, illustrates the view of a "functionning society" held by the implementing agency as communicated to community facilitators, and conveys the view of what constitutes a "successful" society:

Figure 2: Marketing Democratic Practice

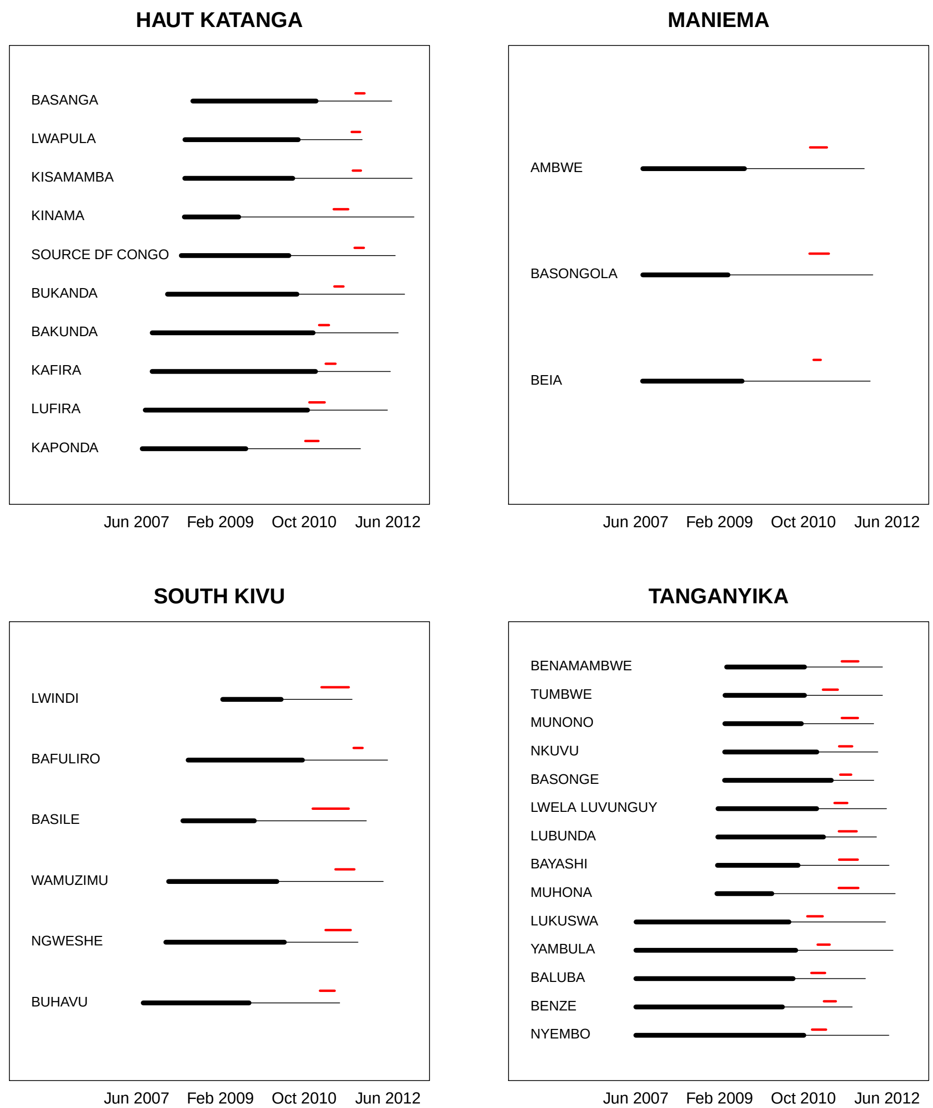


The Tuungane protocols also specified the language to use to market the value of democratic participation:

- La reconstruction dune communauté est laffaire de tout le monde (The reconstruction of a community is everyones business)
- Nous sommes responsables de notre propre destine, reconstruisons notre communaut (We are responsible for our own destiny, lets reconstruct our community)
- Le succès de ce programme dpend rsolument de notre engagement (The success of this program resolutely depends on our dedication)
- Unissons-nous pour le développement (Lets unite for development)
- TUUNGANE, impulser un développement par nous et pour nous, a travers des instances reprsentatives au plus prs de la population (Tuungange calls for development by us and for us, through representation close to the population)
- TUUNGANE, toute la communauté autour dun plan concerte de développement pour vaincre la pauvret et la fracture sociale. (TUUNGANE, the entire community around a development plan to eliminate poverty and social fragmentation.)
- Tout se fait par la communauté et pour la communauté (everything is done by the community and for the community). Source: Protocol 76 - training of CDV.

C Timing of Intervention and Measurement

Figure 3: Timeline of Implementation



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

D Balance

The analyzes in this paper rely on randomization, which guarantees that the treatment and control areas are similar in expectation. In practice, however, it is possible for them to differ simply by virtue of unlucky draws. To test this 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 to the chiefdom capital, presence of infrastructure (specifically: wells, schools, clinics, churches and meeting halls) in 2006, and in-migration in 2006 (IDPs, returned-IDPs, refugees and repatriated refugees). At the individual level we analyze the respondents' age.

Table 5 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. We find that there are no strong differences across the two groups, which is consistent with what is to be expected given the random assignment.

Table 5: Balance

	Control	*Tuungane*	d-stat	N
Distance from major urban center	9.03	8.79	-0.02	802
Presence infrastructure in 2006	7.73	6.94	-0.11	724
In-migration in 2006	7.85	6.92	-0.03	701
Age	39.74	39.3	-0.03	5410

Notes: Based upon the following measures: QE13E, CQ23-27, CQ136-139, QF9. d-statistic is the difference in (weighted) means expressed in (weighted) standard deviations of the control group outcome.

E 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 East 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 East 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?"

F Summary Statistics

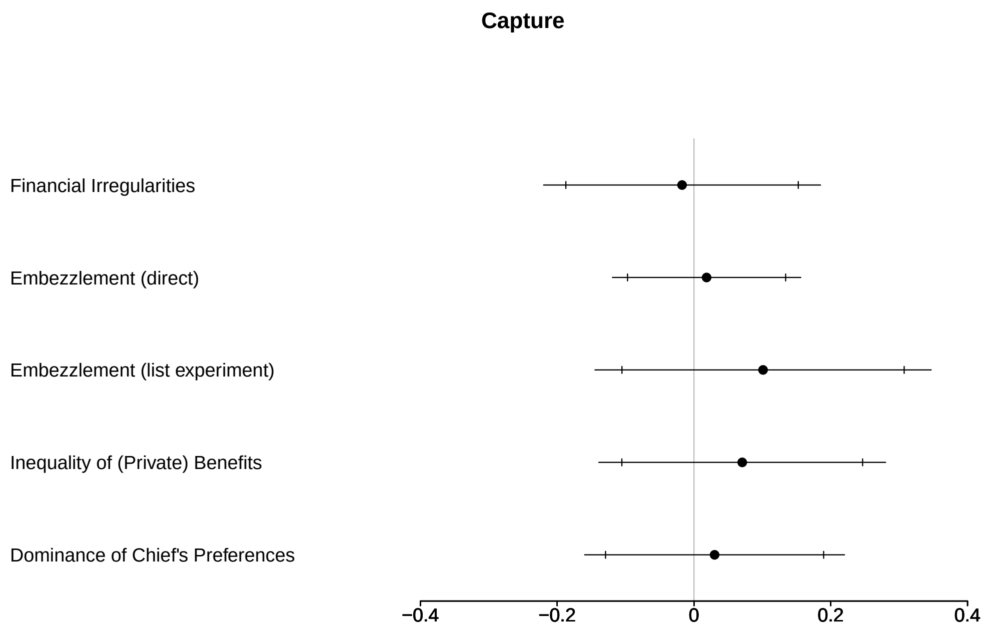
Table 6: Summary Statistics

Variable	RAPID only	Mean Effects	N	Mean	SD	Min.	Max.
Village level variables							
RAPID	NA	NA	1120	0.5	0.5	0	1
TUUNGANE	NA	NA	1120	0.5	0.5	0	1
Financial Irregularities	NA	NA	1120	0.5	0.5	0	1
Embezzlement (direct)	NA	NA	1120	0.5	0.5	0	1
Embezzlement (list experiment)	Y	N	394	0.15	0.21	0	1
Inequality of (Private) Benefits	Y	N	412	0.15	0.22	0	1
Dominance of Chief's Preferences	Y	N	361	0.18	0.73	-2	3
Meeting Attendance	Y	N	409	2.74	5.9	0	35.84
Interventions in Meeting	Y	N	379	-0.08	0.41	-1	1
Dominance of Men in Discussion	Y	N	455	131.13	79.68	20	508
Participatory Selection Methods	Y	N	457	14.54	5.52	1	60
Committee Composition	Y	N	457	73.13	14.66	0	100
Accountability Mechanisms	Y	Y	451	0.06	0.99	-1.49	1.24
Private Complaints	Y	Y	452	0.07	0.96	-2.67	2.04
Knowledge of Project Amount	Y	Y	413	0.04	1.1	-2.26	3
Willingness to Seek Information	Y	Y	412	0.05	0.69	-0.84	2.27
Quality of Accounting	Y	N	411	38.54	28.09	0	100
Information Transmission	N	N	779	38.81	39.48	0	100

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 us 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).

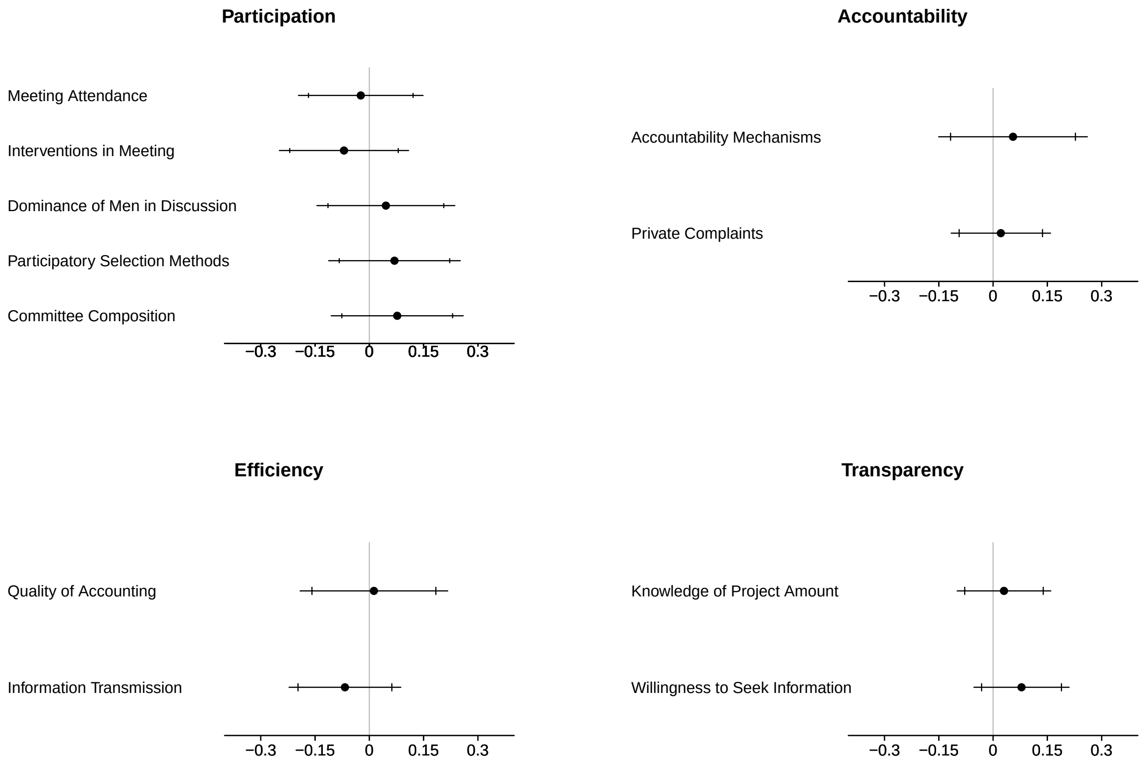
G Graphical Representation of Main Results

Figure 4: Main Results: Graphical Representation



Notes: All analyses employ propensity score weights and clustering of standard errors at the level of randomization clusters. For this figure, all estimates are reported in units of standard deviations of the outcome in the control group. Table 3 is the table version of this figure.

Figure 5: Results on Mechanisms: Graphical Representation

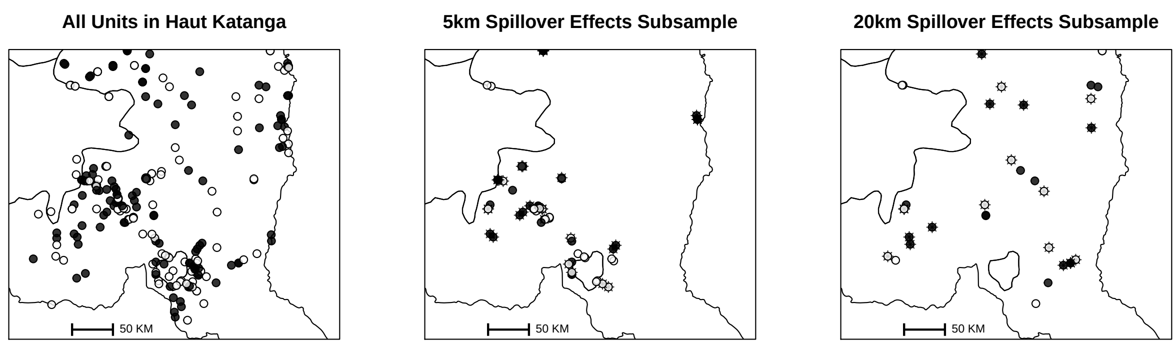


Notes: Treatment effects reported by family. All analyses employ propensity score weights and clustering of standard errors at the level of randomization clusters. For this figure, all estimates are reported in units of standard deviations of the outcome in the control group. Table 4 is the table version of this figure.

H Spillovers

Figure 6 gives an illustration of areas that are included and excluded from the spillover analysis. Units are included only when they have a non 0 probability of being in all combinations of direct and indirect treatment conditions.

Figure 6: 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. Note that here 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 7 and 8 give the results of the spillover analysis that uses inverse propensity score weights and randomization inference.

Table 7: Spillovers at 5km

	Direct	(se)	Indirect	(se)	MSE	(p)	N
Spillovers at 5km							
Financial Irregularities	0.16	0.06	0.01	0.05	0.13	0.89	128
Embezzlement (direct)	0.09	0.05	-0.02	0.05	0.1	0.66	130
Embezzlement (list experiment)	0.27	0.19	0.15	0.15	1.07	0.41	116
Inequality of (Private) Benefits	-0.61	1.23	1.1	1.15	68.39	0.79	130
Dominance of Chief's Preferences	-0.05	0.09	-0.05	0.09	0.38	0.47	124
Participation							
Meeting Attendance	-23.01	12.25	-6.84	11.34	6406.12	0.04	134
Interventions in Meeting	0.74	1.34	1.04	1.24	77.12	0.44	135
Dominance of Men in Discussion	-0.76	3.77	-0.67	3.48	609.09	0.58	135
Participatory Selection Methods	0.26	0.22	0.15	0.2	1.87	0.77	133
Committee Composition	0.22	0.2	0.03	0.18	1.58	0.37	132
Accountability							
Accountability Mechanisms	0.02	0.25	-0.18	0.23	2.73	0.74	132
Private Complaints	0.25	0.17	0.06	0.16	1.25	0.77	130
Transparency							
Knowledge of Project Amount	3.5	5.61	3.56	5.24	1413.94	0.7	130
Willingness to Seek Information	0.37	6.07	-3.75	5.81	4206.16	0.72	263
Efficiency							
Quality of Accounting	-0.37	0.2	-0.34	0.19	1.75	0.63	129
Information Transmission	2.14	1.74	2.03	1.66	353.88	0.22	272

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 maesures are normalized to have zero means. Average direct and indirect effects are then given by α and β. MSE is used as a test statistic for the randomization infernence and the p value reports the probability of such a low MSE under the sharp null of no effects.

Table 8: Spillovers at 20km

	Direct	(se)	Indirect	(se)	MSE	(p)	N
Spillovers at 20km							
Financial Irregularities	-0.01	0.03	0.03	0.03	0.05	0.14	111
Embezzlement (direct)	0.07	0.04	-0.01	0.04	0.08	0.85	121
Embezzlement (list experiment)	0.21	0.14	0.11	0.15	1	0.2	106
Inequality of (Private) Benefits	0.75	0.49	-0.63	0.53	15.75	0.54	121
Dominance of Chief's Preferences	0.1	0.09	-0.06	0.09	0.46	0.27	115
Participation							
Meeting Attendance	4.56	11.99	-28.11	12.49	10470.82	0.96	134
Interventions in Meeting	0.29	1	-0.27	1.04	72.09	0.33	134
Dominance of Men in Discussion	1.32	2.04	2.42	2.13	304.05	0.48	134
Participatory Selection Methods	-0.02	0.14	0.29	0.15	1.52	0.31	135
Committee Composition	0.25	0.14	0.26	0.15	1.52	0.7	136
Accountability							
Accountability Mechanisms	-0.05	0.19	0.2	0.2	2.28	0.33	121
Private Complaints	0.02	0.09	0.02	0.1	0.53	0.57	121
Transparency							
Knowledge of Project Amount	-3.27	4.53	2.96	4.87	1346.38	0.65	121
Willingness to Seek Information	-0.25	5	7.06	5.09	3362.51	0.42	238
Efficiency							
Quality of Accounting	0.02	0.15	-0.09	0.16	1.41	0.15	112
Information Transmission	-0.55	1.84	-0.95	1.91	491.5	0.72	251

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 maesures are normalized to have zero means. Average direct and indirect effects are then given by α and β. MSE is used as a test statistic for the randomization infernence and the p value reports the probability of such a low MSE under the sharp null of no effects.

I Differential Desirability

Table 9: Social Desirability Test

	Positive prompt	Negative prompt	Difference	(se)
Control	0.641	0.834	0.193	(0.025)**
Tuungane	0.633	0.85	0.217	(0.027)**
Difference	-0.008	0.016	0.025	
(se)	(0.03)	(0.022)	(0.037)	

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?" $*p \leq 0.10, **p \leq 0.05, ***p \leq 0.01$.

J 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 7 we found that the implemented 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. This may be due to chief dominance. To explore whether the chief captured the RAPID process in *Tuungane* areas, we investigate whether in *Tuungane* areas, members of the RAPID committee are closer 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.[35] 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.[36] 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 procedures in *Tuungane* areas. During Step B our enumeration team led focus groups with ordinary villagers to learn whether the process of committee and 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 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 areas. We thus conclude that the null result reported in this paper does not reflect chiefs' response to *Tuungane*.

[35]The 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%.

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

K A Bad Case?

One possible explanation for weak effects is that this was a weak intervention.

Did development funders and implementers supporting this project expect 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 project implementers and project directors (12 respondents) as well as a (convenience) sample of seven researchers working in East 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 project 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 projects 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 because it was formed prior to data gathering. Our information on priors supports the idea that the lessons may extend to model cases. Overall, prior beliefs reflect confidence that CDD is an effective model.

Was the project badly implemented? To answer this question we collected data using interviews with members of the population, village chiefs and VDC members in all *Tuungane* villages. The data support the view that the project was well implemented. We find that 62% of the population in *Tuungane* villages knew about *Tuungane* and 39% of those knew who implemented it. Furthermore, 76% of VDC 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 grant amount. We also recorded attendance at project meetings. We find that 35% of the population reported attending some meetings associated with *Tuungane*. More than half of the chiefs interviewed reported attending some meetings and 84% of VDC members reported attendance. The median participant villager attended two meetings, with the top 25% claiming to have attended more than 4. The median participant chief reported attending four meetings while the top 25% attended seven or more; the equivalent numbers for the VDC members are 9 and 15 meetings. The overall knowledge and participation in the project was therefore considerable, and thus the lack of exposure to the democratization components is unlikely to explain our null results.

L 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.

L.1 Attrition and Missing Responses

A first threat to validity stems from missing responses. The study was designed to gather survey data in a sample of 1,120 villages, half of which were selected for the RAPID project. Different targets were set for different items but the most common data (the household survey) was to be gathered for 10 households in RAPID villages and 5 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) data on 72% of villages and 62% of individuals, with higher numbers gathered for steps A and B. The full complement of targeted data was not gathered for a number of reasons.

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 89% of RAPID villages and 89% of 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).[37] This loss covered entire lottery bin areas, 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, we do not think that this loss is plausibly related to the treatment status of units and is thus unlikely to induce bias.

A second significant source of missing data is the inaccessibility of some regions for safety and security reasons. Such losses account for 36 village losses outside Maniema, with balance between RAPID and survey-only villages. However, since these also affect clusters of regions containing both treatment and control areas in nearly equal amounts, they are not plausibly related to treatment status.

A third reason for data missingness is failures in the field, ranging from loss, damage, or theft of PDAs, water damage to paper surveys, or enumerator error in the implementation

[37] A total of 62/147 RAPID villages received Step A, a total 7/147 RAPID villages received Step D. The same number of survey-only villages received Step D.

of surveys or particular questions. Given the difficulty of the environment in the DRC, this third category is relatively small affecting a total of 7% of surveys in surveyed villages. This loss is statistically unrelated to treatment status. The fourth area of data loss is due to non-response on particular questions by subjects, again here we have not found evidence that missingness is associated with treatment status.

A final concern is survey non-response. An examination of household survey data suggests that there was no response from 2,200 out of the 5,473 of the initial households selected for the endline survey; these were replaced by neighboring houses. The major reasons for nonresponse were absence of an individual of the indicated gender (712), empty households (617), refusal for any member of the household to be interviewed (95, or <2%), and households that were not found by the survey teams (360).[38] These non-responding households were split almost exactly half and half between treatment and control units suggesting that household missingness is not correlated with treatment. The implication of this is that household level results can only be interpreted as the attitudes of individuals in households accessible at time of survey.[39]

L.2 Noncompliance

A second threat to the validity of the interpretations offered here is treatment noncompliance in the sense that areas that were selected by lottery to participate in *Tuungane* did not, and areas that were not selected did in fact participate. 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 project 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

[38] In Step A enumerator teams created a sampling frame of all households in the village. From this ten households were randomly selected: five to be interviewed during Step A and Step D, and five to be interviewed only during Step D. For non-RAPID villages the sampling frame was created, and five households selected, during Step D.

[39] And more precisely of accessible individuals in accessible households.

is no failure of compliance. Our results (see Table 10) are similar.

L.3 Treatment Heterogeneity

As seen in Figure 3 there is heterogeneity both in the timing and length of project implementation and the timing and length of data collection relative to project 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 project 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 bins, all units in lottery bin areas were first exposed to the project at the same time (although projects 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 bin 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.

L.4 Specification

Out of concern that analysis decisions resulted in false negatives, 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 control for lottery bin fixed effects. Finally, 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 10) are robust to these different specifications.

L.5 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, Reed, and Robinson, 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 11 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: Robustness

	Base	Alt. Treat.	Alt. Specifications Village	Bins	Prop.
	(se)	(se)	(se)	(se)	(se)
Financial Irregularities	-0.004	0.005	-0.004	-0.006	-0.001
	(0.021)	(0.021)	(0.021)	(0.02)	(0.021)
Embezzlement (direct)	0.006	-0.01	0.006	0.014	0.003
	(0.024)	(0.024)	(0.021)	(0.02)	(0.022)
Embezzlement (list experiment)	0.107	0.096	0.078	0.144	0.079
	(0.133)	(0.133)	(0.077)	(0.117)	(0.076)
Inequality of (Private) Benefits	0.391	0.277	0.196	0.253	0.201
	(0.593)	(0.592)	(0.583)	(0.524)	(0.591)
Dominance of Chief's Preferences	0.013	0.002	0.009	0.011	0.01
	(0.043)	(0.043)	(0.042)	(0.044)	(0.042)
Meeting Attendance	-1.983	-2.731	-1.983	-1.199	-1.556
	(7.367)	(7.367)	(7.367)	(6.281)	(7.466)
Interventions in Meeting	-0.391	-0.252	-0.391	-0.391	-0.167
	(0.509)	(0.509)	(0.509)	(0.509)	(0.508)
Dominance of Men in Discussion	0.651	0.137	0.651	0.612	0.579
	(1.385)	(1.385)	(1.385)	(1.279)	(1.374)
Participatory Selection Methods	0.07	0.08	0.07	0.069	0.076
	(0.093)	(0.093)	(0.093)	(0.073)	(0.093)
Committee Composition	0.076	0.083	0.076	0.092	0.069
	(0.091)	(0.092)	(0.091)	(0.075)	(0.09)
Accountability Mechanisms	0.056	0.02	0.056	0.054	0.025
	(0.107)	(0.107)	(0.107)	(0.103)	(0.107)
Private Complaints	0.022	-0.004	0.01	0.004	-0.001
	(0.072)	(0.073)	(0.068)	(0.052)	(0.069)
Knowledge of Project Amount	1.462	1.33	1.288	2.29	1.791
	(3.207)	(3.211)	(2.782)	(2.541)	(2.8)
Willingness to Seek Information	3.839	3.753	3.661	3.058	2.917
	(3.277)	(3.223)	(2.827)	(2.265)	(2.843)
Quality of Accounting	0.013	-0.012	0.013	0.011	0.006
	(0.105)	(0.105)	(0.105)	(0.084)	(0.105)
Information Transmission	-1.339	-0.738	-0.461	-1.476	-0.873
	(1.565)	(1.622)	(1.466)	(1.53)	(1.457)

Notes: 'Base' corresponds to the results reported in Table 3. "Alt. Treat." are results using a treatment variable that uses IRC's classification of treatment in cases in which databases disagreed. "Village" are results in which all variables are aggregated to the village level using individual sampling weights. "Bin" are results at the village level introducing controlling for lottery bins. "Prop." are results (at the village level) using propensity weights adjusted to assess village level sample average treatment effects rather than population average treatment effects. $*p \leq 0.10, **p \leq 0.05, ***p \leq 0.01$.

Table 11: Heterogeneous effects of grassroots democratization, by initial institutions

	Base	Projects	Committees	Inherited
	(se)	(se)	(se)	(se)
Financial Irregularities	-0.004	0.005	0.004	-0.047
	(0.021)	(0.031)	(0.042)	(0.034)
Embezzlement (direct)	0.006	0.013	0.01	0.003
	(0.024)	(0.035)	(0.038)	(0.03)
Embezzlement (list experiment)	0.107	-0.045	0.608	0.3
	(0.133)	(0.203)	(0.219**)	(0.22)
Inequality of (Private) Benefits	0.391	1.542	0.509	1.858
	(0.593)	(0.963)	(1.126)	(0.994)
Dominance of Chief's Preferences	0.013	0.018	0.056	-0.014
	(0.043)	(0.064)	(0.069)	(0.07)
Meeting Attendance	-1.983	-4.956	-9.803	-1.248
	(7.367)	(11.034)	(15.164)	(12.761)
Interventions in Meeting	-0.391	-1.072	-1.233	-0.358
	(0.509)	(0.951)	(1.169)	(0.915)
Dominance of Men in Discussion	0.651	-1.549	1.354	0.163
	(1.385)	(2.318)	(2.727)	(2.23)
Participatory Selection Methods	0.07	0.326	0.106	0.389
	(0.093)	(0.152**)	(0.166)	(0.16**)
Committee Composition	0.076	0.072	-0.161	0
	(0.091)	(0.155)	(0.192)	(0.158)
Accountability Mechanisms	0.056	-0.116	-0.198	0.027
	(0.107)	(0.163)	(0.195)	(0.178)
Private Complaints	0.022	0.087	-0.043	0.149
	(0.072)	(0.105)	(0.123)	(0.106)
Knowledge of Project Amount	1.462	2.579	2.074	0.52
	(3.207)	(4.793)	(5.586)	(4.938)
Willingness to Seek Information	3.839	8.429	1.406	2.658
	(3.277)	(4.818)	(5.842)	(4.678)
Quality of Accounting	0.013	-0.065	0.174	-0.075
	(0.105)	(0.165)	(0.19)	(0.187)
Information Transmission	-1.339	-4.446	-1.375	-1.562
	(1.565)	(2.835)	(2.945)	(2.512)

Notes: This table replicates Table 3 for three subgroups. The column under 'Base' are the results as presented in Table 3. 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. $*p \leq 0.10, **p \leq 0.05, ***p \leq 0.01$.

M Registration and Mock Report

Conscious of concerns that empirical analyses can suffer from a propensity to favor reporting "significant" findings in classical statistical tests, and that this practice can lead to bias in assessment of effects, we sought to employ a form of pre-registration of our research design. All of our analysis 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– circulated to colleagues and posted online (this was prior to the existence of a social science registry to which we could post the analysis plan). The analysis presented here differs from those described in the analysis plan in four ways.

First, we focus here 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, Sanchez de la Sierra, and Van der Windt (2012).

Second, we made five adjustments to outcomes related to capture. Financial irregularities is presented as the share of the $1,000 instead of the dollar amount. 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). We also chose not to condition the latter by those villages that chose distribution projects. We added the two embezzlement measures. 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. The results are unaffected.

Third, the index on health information flows was changed to focus only on items that were provided to peers (excluding items provided uniquely to chiefs). This was to reflect the intention of the original measure but produces no substantive effect on results.

Fourth, for a few tables, we added summary analyses, as described in Section 6.3. These make for easier interpretation of the multiple results described in the tables.

Finally, the registration document and mock report covered the main hypotheses and tests; the tests provided in Sections L and 8.2.2 were not elaborated in the mock report.

These deviations do not alter the results. All results are available as originally specified in Humphreys, Sanchez de la Sierra, and Van der Windt (2012).