

Do Selection Rules Affect Leader Responsiveness? Evidence from Rural Uganda

Guy Grossman*

University of Pennsylvania, Philadelphia, PA, USA; ggros@sas.upenn.edu

ABSTRACT

Community organizations in developing countries often suffer from self-serving elites. This study examines whether the responsiveness and accountability of local leaders can be strengthened through the introduction of more inclusive and participatory leader selection rules. To address identification problems, I take advantage of natural conditions that resulted in exogenous variation in the rules for selecting leaders of farmer associations in Uganda. I find that compared to leaders appointed by the community elites, directly elected leaders are

* I am grateful to Sylvie Hoster, Alex Barnard, Eliana Horn, Vivian Lu, and my Ugandan research team for their excellent research assistance; to Macartan Humphreys, Tim Frye, Kimuli Kasara, Alex Scacco, Jordan Kyle, Elizabeth Sperber, Eric Mvukiyehe, Kate Baldwin, Thomas Zeitzoff, and Walker Hanlon for their helpful comments on earlier drafts. The paper has also benefited tremendously from participants at CAPERS 6, EPSA 2011, and MPSA 2012 and at seminars and workshops at Yale, Stanford and Madison Wisconsin. I gratefully acknowledge support from the NSF Doctoral Dissertation Improvement Grant SES-0921204. This manuscript had previously circulated entitled “Causal effects of leader selection rules on leader responsiveness and cooperation: evidence from Ugandan community organizations.”

Online Appendix available from:

http://dx.doi.org/10.1561/100.00013012_app

Supplementary Material available from:

http://dx.doi.org/10.1561/100.00013012_supp

MS submitted 19 January 2013; final version received 29 July 2013

ISSN 1554-0626; DOI 10.1561/100.00013012

© 2014 G. Grossman

significantly more responsive to group members, leading to greater cooperative behavior. Analyzing possible mechanisms, I find that community organizations using appointments are less likely to develop monitoring institutions that are vital for constraining the behavior of local elites. Unique social network data provides evidence that close friendship ties between appointed and appointers substitute for formal monitoring institutions, leading to loss of confidence by community members and, subsequently, to a decline in public goods contributions.

Keywords: Political accountability; political selection; community-driven development; monitoring; local public goods.

In many low-income countries, the public sector is failing to provide adequate levels of public goods and social services. In response, recent years have witnessed the proliferation of *non-state*, community-level collective action initiatives. Supported by the international aid community, as part of a paradigm that stresses the positive effects of participatory, community-driven development, communities are increasingly assuming responsibility over maintaining local infrastructure projects (Khwaja, 2009), selecting and executing development programs (Casey *et al.*, 2012), raising funds to support public schools (Miguel and Gugerty, 2005), and monitoring health clinics (Bjorkman and Svensson, 2009). In addition, various community organizations, such as women, farmer and credit cooperatives, provide local services on a membership basis (Gugerty and Kremer, 2008).¹

Since the delivery of local public goods depends, to a large extent, on the voluntary contributions of community members, the effectiveness of community organizations crucially depends on ensuring that the benefits of development are not captured by local elites (Bardhan and Mookherjee, 2006).² The extent to which *formal governance institutions* — for example, participatory decision-making procedures or the method for selecting local leaders — can constrain the behavior of local leaders is still an open

¹ Empirical evidence documents the proliferation of community organizations: in Senegal 10% of sampled villages reported having at least one such group in 1982; by 2002 this figure was 65%. In Burkina Faso the figures were 22% for 1982 and 91% in 2002 (Bernard *et al.*, 2008).

² Beyond weak accountability, variation in the quality of local public goods is also explained as a function of communities' social capital (Krishna, 2002), and other community characteristics, such as size, income distribution, or level of ethnic homogeneity (Alesina and La Ferrara, 2005).

question. Specifically, leading scholars argue that attempts by outsiders to improve existing governance institutions are likely to be futile since historically rooted formal and informal rules and social norms are rather difficult to transform (Easterly, 2006).

This paper contributes to the current debate by examining the causal effects of introducing participatory leader selection processes at the community level. To overcome identification challenges inherent to the study of political institutions, I take advantage of natural conditions that resulted in plausibly exogenous variation in the rules for selecting the leaders of Ugandan farmer associations. Created around 2006 as part of one of Uganda's largest recent rural development programs, all the farmer associations in the study area share a similar organizational structure. However, in approximately half of the farmer associations, community members vote directly for their preferred candidate, whereas in the remaining associations the leader is appointed by a small number of *elected representatives* serving on the organization's council.

The paper's identification strategy — developed more fully below — rests upon the observation that (a) project field-trainers, experienced extension agriculture officers hired and trained centrally by the NGO implementing the program, played a key role in the process of group formation; (b) almost all farmer organizations followed their trainer's recommendation when adopting a leader selection rule; (c) the deployment of field-trainers to various districts was “as good as random;” (d) trainers' recommendations were based on personal preferences that were unrelated to characteristics of the groups; and (e) the facilitation process did not vary significantly, apart from the field trainers' recommendation of a leader selection rule.

Using the field-trainers' recommendation as an instrumental variable, I find that associations using direct elections result in leaders who, on average, are more responsive and accountable to community members. I then test the proposition, suggested in various cross-national studies and laboratory experiments, that strengthening leader accountability leads to greater willingness to contribute toward the provision of local public goods. I find that direct elections increase members' cooperative behavior, resulting in the provision of a wider range of group services.

Turning to the analysis of causal mechanisms, I demonstrate that monitoring institutions are a key mediator in the relationship between leader selection rules and leader responsiveness. Consistent with a theoretical framework developed below, I find that community organizations using

appointments are less likely to invest in monitoring institutions that are necessary for constraining the behavior of local elites. Using unique social network data I provide evidence that close friendship ties between appointed and appointers substitute for formal monitoring institutions, leading to loss of confidence by community members and, subsequently, to a decline in public goods contributions. Using a set of behavioral experiments (dictator and third-party-punishment games) I also find evidence suggesting that directly elected leaders are more committed to the welfare of community members. Contrary to my initial expectations, I do not find evidence that direct elections are associated with greater vertical accountability; that is, bottom-up grassroots participation.

The findings of this paper complement and expand recent work on the impact of governance institutions in the context of Community Driven Development (CDD) programs in three important ways. A central objective of CDD studies is to assess the relative effectiveness of rules for selecting and implementing *development projects*. This paper focuses, instead, on the impact of the method for selecting *local leaders*. Second, CDD studies focus on community-wide organizations. Conversely, this study considers governance institutions in *voluntary* associations in which members have a more credible exit strategy. Third, the paper uses complete network data of community elites in all (fifty) localities, not available to previous studies. These data enable me to better identify the mechanisms that link between formal governance institutions and elites' behavior.

The paper's findings are also relevant to a literature on the effect of introducing elections in non-democratic settings. Baldwin and Mvukiyehe (2011) argue that the introduction of elections in place of traditional forms of appointments in Liberia resulted in clan chiefs that are more likely to articulate goals that are aligned with community members. Similarly, using variation in the timing of top-down introduction of elections in Chinese villages, Martinez-Bravo *et al.* (2011) and Luo *et al.* (2007) argue that elected village heads are more responsive to the priorities of community members, compared to party appointees. In a recent experiment, Beath *et al.* (2013b) find that, under certain conditions, elected village councils in Afghanistan are better at targeting development aid than traditional elites. Though valuable and informative, since those studies compare between democratic elections and non-democratic selection methods — appointment by unelected party officials in China or by traditional leaders in Liberia and Afghanistan — they have a hard time separating the effect of the political institution from the

effect of the perceived legitimacy of the office-holder. This paper is better positioned at isolating the effect of the political process, since it juxtaposes equally legitimate democratic selection rules.³

Finally, the paper's findings are also relevant to a political economy literature that debates the tradeoffs between election and appointment of public officials across US states. Examples include judges (Hanssen, 1999), heads of regulatory agencies (Besley and Coate, 2003) and members of school boards (Hess, 2008) who are appointed by politicians in some jurisdictions, but are directly elected by the public in others. It is hard, however, to draw definitive conclusions from this literature mainly because political units endogenously adopt rules for selecting officials (See critique by Cavazos (2003)). In addition, since its empirical evidence relies on aggregated state-level cross-sectional data this body of work has limited ability to uncover micro-foundational processes associated with different selection methods.

1 The Debate Regarding Formal Governance Institutions

This paper most directly addresses the ensuing debate regarding the effectiveness of *formal* governance institutions in constraining the behavior of local leaders in low-income countries. On one hand, it is argued that the introduction of inclusive and participatory rules increases the responsiveness of local leaders and reduces elite capture (Platteau and Abraham, 2002). This position follows arguments by institutional design theorists recommending devolving power to local levels together with introducing democratic institutions so that citizens can monitor and sanction officials effectively (Seabright, 1996). Empirical evidence includes, among others, studies by Fritzen (2007) and Dasgupta and Beard (2007) in Indonesia, showing that community leaders' rent-seeking behavior is reduced by introducing democratic arrangements. Similarly, Labonne and Chase (2009) show that participatory rules for selecting development projects in the Philippines resulted in equal representation of non-elite preferences.

On the other hand, leading scholars question whether institutional design can substantially affect the behavior of local leaders, especially where traditional hierarchies and economic inequalities likely make informal rules resilient to change. Indeed, political institutions in developing countries,

³ There are some concerns regarding the identification strategy of Martinez-Bravo *et al.* (2011) and Luo *et al.* (2007): Chinese provincial governments likely tested the impact of elections in carefully chosen villages.

particularly those inherited from the colonial period, are thought to be subject to strong inertia (Acemoglu, Johnson and Robinson, 2001), and thus difficult to transform in the short-term (Easterly, 2006). Bowles and Gintis (2004), argue that formal institutions will continue to have a limited impact at the community level, so long as mechanisms of democratic accountability are weak. Thus, to be effective, argues Bardhan (2002), institutional reforms must be accompanied by changes of existing power structures. According to this view, newly introduced formal governance institutions can constrain the behavior of leaders only in communities *that are already relatively equal and egalitarian*. Findings from a recent CDD study by Araujo *et al.* (2008) that elite domination is greater in communities with greater economic inequality, is consistent with this position.

Addressing this debate has more at stake than simply resolving academic quibbles. Local public goods provision is an important development practice and community programs are a primary mean by which aid is currently being disbursed.⁴ The current debate suffers, however, from several limitations. First, the existence of inclusive, participatory rules are likely to be correlated with unobserved features of a community that may independently determine leader accountability. Second, high levels of participation in decision-making processes may cause improved project performance, but improved project performance often also encourages greater grassroots participation. Unfortunately, most previous work had limited success in dealing with underlying identification issues inherent in the highly endogenous structures of political institutions (Mansuri and Rao, 2013).

Three recent field experiments of CDD program — in Liberia (Fearon *et al.*, 2009), Sierra Leone (Casey *et al.*, 2012), and the Democratic Republic of Congo (Humphreys *et al.*, 2013) — were designed explicitly to contribute to the current debate. CDD projects invest resources in building local governance capabilities before distributing development funds to the community in the form of block grants. If the introduction of new formal governance institutions makes local leaders more accountable, then treatment villages are expected to have an advantage in mobilizing resources from community members. In both Sierra Leone and the DRC, however, the authors find no evidence that the program led to fundamental changes in the ability to raise funds for local public goods. By contrast in Liberia, the authors find

⁴ The World Bank alone currently spends about \$2 billion per year on various CDD programs, according to the World Bank's CDD website.

that treatment communities contributed significantly more in a public good game, compared to control communities. Specifically, leaders in Liberian communities in which democratic governance rules were introduced engaged in greater mobilization and information-sharing efforts that increase coordination around socially desirable outcomes.

Reconciling these differences is hard given that CDD programs are, in effect, composite interventions that include many components (e.g., institutional building, civic education, project funding, in-kind matching contributions, and financial management). This bundling is limiting the ability of the recent experiments to separate between the impact of the newly introduced governance rules and other aspects of the program.⁵ One key advantage of this study is that since the only difference between “treatment” and “control” communities is the leader selection rule, it is better positioned than past studies to identify the impact of the political institution itself.

Finally, two recent field experiments have attempted to isolate the impact of political institutions, in a CDD context, by randomizing the rules for selecting *development projects* (Olken, 2010; Beath *et al.*, 2013a). To the best of my knowledge, no study has yet to successfully isolate, instead, the rules for selecting *community leaders*. This is understandable given the various political and ethical issues surrounding randomizing the rules for selecting local leaders and the difficulty of identifying natural conditions resulting in exogenous variation. By examining the impact of leader selection rules, this paper broadens the range of studied governance institutions that may affect incumbents’ behavior at the local level.

2 Theoretical Argument

In this section, I provide a simple theoretical account for why directly elected community leaders are likely to be more responsive and accountable than appointed office-holders. I begin by exploring the relationship between leader

⁵ Notwithstanding the broad similarities between programs, there are also important differences in program design between the three field experiments that might be consequential. Those include, but are not limited to, the definition of communities, the planned timeline, the identity of the implementing NGO, the intensity of intervention, the number of intended beneficiaries, the size of the block grants, the list of projects and activities that cannot be funded via the block grants, whether and how much the community itself contributes to financing, the rules that govern project selection, whether the program is implemented with new or existing community institutions, etc.

selection rules and *monitoring institutions*, before tuning to explore their relationship with *leader type*.

2.1 *Leader Selection Rules and Monitoring Institutions*

The closer local leaders are monitored the more likely they are to be responsive to their constituents. I argue that directly elected leaders are more constrained in their behavior because direct elections likely result in higher levels of monitoring. There are two mechanisms that tie between leader selection rules and communities' monitoring level: (a) vertical accountability and (b) horizontal accountability mechanisms.

Vertical accountability denotes here enhanced grassroots participation by community members. By making community members responsible for holding office-holders to account, direct elections likely incentivize villagers to invest in obtaining information on the behavior of incumbents and on the performance of their community organization, in order to make an informed vote choice. This sort of grassroots participation changes incumbents' behavior as it reduces information asymmetries that underline moral hazard. Consistent with this argument, Bjorkman and Svensson (2009) show that health clinic staff in rural Uganda exert higher efforts in serving their communities in response to grassroots monitoring efforts. In addition, the experience of participating in the selection of leaders may provide a training ground for other types of political involvement (Lindberg, 2006) that constrain the ability leaders to dominate decision-making processes. I formalize these intuitions in the following hypothesis:

H_1 **Vertical accountability:** Community members are more likely to actively participate in local governance under direct elections rule.

Horizontal accountability refers here to the relationship between different office holders. Building on the corporate governance literature, I argue that community organizations using appointments are less likely to develop formal monitoring institutions that serve to constrain the behavior of local leaders. A key insight of the corporate governance literature is that appointment systems face the risk of (growing) familiarity between appointed office-holders and their appointers. Familiarity or friendship is problematic due to possible *substitution effects*: the idea that *formal* monitoring institutions and "*closeness*" between appointed and appointers are seen as substitutes. Such familiarity, therefore, is more likely to induce lax oversight

and a governance culture that makes it easier for office-holders to control data, limits accountability and reduces opportunities for constituents' input (Shaub *et al.*, 2005). From a theoretical standpoint the key insight is that leader selection institutions do not merely affect incumbent-constituents relation, but also decisively shape the relations between office-holders, in our case community elites.

***H*₂ Horizontal accountability:** Community organizations using appointments are less likely to develop formal auditing institutions and monitoring practices.

2.2 Leader Selection Rules and Leader Type

Leaders are expected to be more responsive to the preferences of the median voter of the group that selects them (Bueno de Mesquita *et al.*, 2003). To the extent that the preferences of the small number of appointers — usually members of the village elite — are different from the preferences of the average community member, direct elections should change the orientation of incumbents. This shift in accountability finds some support in empirical studies (Martinez-Bravo *et al.*, 2011), and in normative theories of representation (Mansbridge, 2003).

Against this reasoning, which implicitly assumes retrospective-voting accountability mechanism, Besley and Coate (2003) argue that a finding that directly elected office-holders are more responsive than appointed ones is hardly intuitive. If representative democracy yields median-voter outcomes, then *elected* representatives should have an incentive to appoint office-holders accepted by the majority of the community, producing similar outcomes as direct elections. Accounting for their finding that direct elections of state utility regulators result in more responsive office holders, Besley and Coate (2003) suggest, but do not test, that appointment systems suffer from a bundling effect.

Bundling refers to the idea that in appointment systems constituents may vote for a candidate for reasons other than his or her likelihood of appointing an office-holder that is committed to advancing citizens' preferences. Unless the identity of the appointed official is very salient, elected representatives may not feel bound to appoint office-holders committed to the welfare of the median community member. This sort of bundling is avoided in direct elections, where constituents have one vote for their representative and one for

the office-holder. When public officials are directly elected, their responsiveness to the preferences of constituents likely dominates constituents' vote choice. In sum, Besley and Coate (2003) discussion points to the idea that direct elections allow citizens not simply to hold incumbents to account for past performance, but rather to use their vote to select leaders who are closer to them in preferences, beliefs or some ascriptive characteristic.⁶ Though a direct testing of "bundling" is almost never feasible (Besley and Coate, 2003) we can study the plausibility of its logic by examining the following testable implication:

H_3 Leader Type: Directly elected leaders are more likely to share observable characteristics with the "average" community member.

Closely related, direct elections may increase local leaders' responsiveness simply due to norms of reciprocity. The fact that community members elected the incumbent (rather than his/her challenger) could bring with it greater expectations for prosocial behavior, and there may be reputation costs for not meeting such expectations (Baldassarri and Grossman, 2013). In this case directly elected leaders are more likely to take into consideration members' welfare.

H_4 Reciprocity: Directly elected leaders are likely to exhibit greater commitment to the welfare of community members.

This study was explicitly designed to identify which of the above mechanisms is more likely to be a mediator between leader selection rules and leader accountability.

3 The Research Site: APEP Farmer Associations

I explore the impact of leader selection rules on leader responsiveness in the context of farmer associations, which are a pivotal community organization in many low-income countries. Farmer associations' *raison d'être* is to improve the performance of their members' farms as economic units. Members join voluntarily to gain access to the services produced by the association, of which the most important is securing higher output prices

⁶ This mechanism is close in spirit to the notion *descriptive representation* developed by Pitkin (1967).

through collective marketing. Other services include securing lower input prices and agriculture training.

It is worth considering briefly the history of farmer cooperatives in East Africa. In low-income African countries such as Uganda, the livelihood of the majority of citizens depends on agriculture. According to the 2002 census, about 88% of Ugandans live in areas that are regarded as rural, many of which are subsistence producers who do not participate actively in the monetized economy. Policies for integrating smallholder farmers into markets have, therefore, been rather salient for both colonial and post-colonial governments (Bates, 1981). Yet despite numerous efforts to collectivize production and crop marketing, many past initiatives have failed, to a large extent because colonial and national governments sought to control and exploit farmer organizations (Lele, 1981).

In the 1990s, government-controlled farmer organizations across Africa had become insolvent as a result of structural adjustment reforms (Ponte, 2002). This process triggered an immediate renewal in the interest of the aid community in supporting self-governed farmer organizations (Hussi *et al.*, 1993). In recent years, development agencies have been dramatically increasing their funding for farmer organizations, which are believed to be essential for integrating small-scale producers into markets. Recent studies assessing the contribution of aid-supported farmer groups to raising income and farm output show, however, mixed results (Ashraf *et al.*, 2009).

3.1 Governance Structure of APEP Community Organizations

The community organizations considered herein were created as part of the Agriculture Productivity Enhancement Project (APEP), which is one of the largest recent rural development projects in Uganda.⁷ To achieve its main goal of deepening the transition of small-scale producers into commercial farming, APEP organized about 60,000 farmers into over 2,500 village-level producer organizations (POs) in 30 districts. It soon became apparent that POs lack the capacity and bargaining power to become viable economic entities. In response, APEP began encouraging neighboring POs to form larger cooperatives. The resulting farmer associations (known as Depot Committees or DCs) were organized at the parish-level, typically covering a cluster

⁷ APEP was funded by USAID and implemented by Chemonics International, an International Development consulting firm based in Washington DC. The project was launched in 2004 and ended in July 2009.

of ten nearby villages, approximating natural communities. The DCs were designed to exploit economies of scale, to reduce transaction costs, and to increase the bargaining power of farmers through collective marketing.

Ugandan field-trainers, deployed by the NGO implementing the project, presided over the process of group formation. For this reason, all APEP associations share a similar organizational structure. The association's main governance body is the DC council, comprised of two *elected representatives* (councilors) from each of the embedded producer organizations (POs). The responsibilities of the councilors include, among others, transmitting information from and to their respective POs, participating and voting in plenary sessions and serving on one of the association's standing committees. These include the executive committee as well as loans, marketing and procurement committees.⁸ One of the key responsibilities of council representatives is to monitor the work of the senior executive officer: the DC manager.

The DC manager plays a pivotal role in organizing the farmer groups' main public good: collective marketing. The DC manager's most important responsibilities include searching for buyers, negotiating input and output prices, and organizing the collection of crops (including hiring and supervising employees). Additionally, managers organize training workshops, facilitate the flow of information throughout the association and coordinate sanctioning of members who do not follow the association's rules and bylaws. A graphical representation of the association governance structure is provided in the online appendix. As noted, there exist two selection methods for DC managers across the APEP organizations. In about half of the farmer associations, registered members vote directly for their preferred candidate. In the remaining associations, council representatives, elected by their respective POs, appoint the manager. Below I describe, in detail, the study's strategy for identifying the causal impact of those leader selection rules.

4 Sampling and Data Sources

This section describes the data used in the paper. To reduce crop-related variability, I limited the study's target population to 105 farmer associations

⁸ The executive committee is comprised of the manager, a chairperson, a secretary, and a treasurer. The DC council decides which representative serves on what committee and at what role. Prestigious positions, such as the council chairperson, tend to be contested and decided through some voting procedure. Allocation to other positions depends on the interests, expertise and time constraints of council representatives.

marketing coffee, the most common crop produced by the APEP groups. I then sampled 50 associations from 5 district-areas using a stratified, random, multistage cluster design. Quantitative data for the empirical analysis were collected between July and September 2009 by local interviewers.

Within each association, different types of data were collected. At the association level, a questionnaire was completed via an interview with the four members of the DC executive committee. This questionnaire provided information on the group's governance rules — for example, frequency of council meetings, term lengths and term limits, presence of a constitution — including the method for leader selection. I refer to this data source as the “executives’ survey” or (E). Data on a DCs economic activities were also derived from its books and records. For each association, I sampled six producer organizations (POs), for a total of 287.⁹ Interviews with the leaders of the sampled POs provide additional information at that level. From each sampled PO, I further sampled an average of six members, for a total of 36 members per association.¹⁰ Sampled members were surveyed in person by trained interviewers in the respondents’ local language for a total of 1,781 surveys. I refer to this data source as the “members’ survey” or (M).

In addition, I surveyed the entire set of council representatives, whether or not their PO was selected into the sample, for a total of 1,316 interviews. These “representatives’ surveys” (R) only partially overlap with the members’ surveys, as they were tailored to capture the roles and responsibilities of councilors within the association’s organizational structure. Each surveyed councilor also completed a social network module using a roster of the names of all other council representatives. The network data allowed me to analyze the structure of the DC leadership network (e.g., its density), as well as the network position (e.g., centrality) of each councilor. The survey team visited each community up to four times to reduce attrition, which was reduced to about 10%.

5 Plausibly Exogenous Variation in Governance Rules

The paper’s identification strategy rests upon three pillars. First, project field-trainers played a key role in establishing the organizational structure of the APEP groups. Second, the newly created groups almost always adopted

⁹ In a few cases, where a DC had fewer than seven POs, all were selected.

¹⁰ The number of sampled members from each sampled PO was proportional to the size of the PO, to assure that the sample is self-weighted.

their trainer’s recommendation when choosing governance institutions such as a leader selection method. Third, trainers’ recommendations were orthogonal to group characteristics. I now turn to evaluate the plausibility of these assumptions.

5.1 *The Instrument*

APEP field-trainers helped village representatives to agree on the structure and rules that govern their farmer association. This process was informed by a manual that outlined the steps field-trainers should take to establish a DC. The centrality of the trainers in the process of group formation is attested to by the fact that all APEP associations, regardless of the trainer’s identity, have a similar organizational structure and power division between the POs and the DC.

The manual did not address, however, every aspect of group formation. Importantly, it did not specify explicitly the selection rule for the DC manager position.¹¹ Five of the eight trainers in the study area encouraged the farmer groups to endow the council representatives with the power to appoint the DC manager. The remaining field-trainers recommended, instead, that group members vote directly for their preferred candidate for the manager position. In personal interviews, field-trainers explained that their recommendation was based on what *they* considered to be the “most appropriate” selection method.¹² With this in mind, I propose to use the trainers’ recommendation as an instrumental variable (IV) for a group’s leader selection rule.

5.2 *Assessing the Validity of the Instrument*

The counterfactual that forms the basis of the paper’s identification strategy is the change in the outcomes of interest had APEP deployed to community organization j , a field-trainer with a different leader selection rule

¹¹ In interviews I conducted with the APEP management they referred to the fact that the manual was not more explicit about some governance rules as an “error.” A copy of the manual can be shared upon request.

¹² This error is likely due to the fact that trainers recommending appointments considered a selection by the DC councilors as an election procedure, and because APEP trainers were trained as agriculture extension officers and were, therefore, not accustomed to think about possible tradeoffs embedded in various leader selection methods.

preference.¹³ Out of 50 farmer associations that were randomly sampled, 32 were encouraged to use appointments, of which 25 associations (78%) complied. Similarly 17 of the 18 associations (94%) that were recommended to use direct elections adopted that rule (see online appendix, Table A2). These high take-up rates further attest to the centrality of the field-trainers in the resulting choice of governance institutions.

Following the literature on encouragement research design (Angrist and Pischke, 2009), I use the field-trainers' recommendations as an instrumental variable. This allows me to identify an unbiased local average treatment effect (LATE) for "compliers"; farmer associations that used direct elections, but would have used an appointment system had their trainer made a different recommendation. Encouragement designs require satisfying several conditions. First, the relationship between treatment assignment (the instrument) and the endogenous treatment variable, the actual leader selection method, must be strong. First-stage regression of leader selection rule on the trainer's recommendation produces a positive and significant coefficient (p -value = 0.00).¹⁴

Second, the instrument needs to be independent of potential outcomes. There are several reasons to believe that the instrument satisfies the exogeneity condition. First, the deployment of trainers to project areas was "as good as random," since it was based on criteria that were *unrelated to group attributes*.¹⁵ Second, as noted, trainers have based their recommendation on their beliefs regarding the appropriate selection method, that is, not on groups' attributes. Had trainers conditioned their recommendation on group characteristics, variability should be witnessed in their recommendations. Trainers, however, made the *same* recommendation to *all* the DCs they helped create, including those that are out-of-sample (online appendix, Table A2). The fact that *different* recommendations were made

¹³ The study's main results are robust to an alternative counterfactual that operates at the district-level. One-sided randomization inference p -values of the main summary index are significant when the randomization inference analysis imposes the same leader selection rule on all DCs within the same district.

¹⁴ F statistics for joint significance of the instrument in first-stage regression far exceed the widely used critical value of 10. There are no reasons to suspect that monotonicity and homogenous partial effects (Dunning, 2008) — the fourth and fifth conditions — do not hold in this case.

¹⁵ The first criterion guiding trainers' group assignment was APEP's decision to deploy trainers to areas *other than their home district*. The rationale was to reduce the likelihood that trainers engage in income-generating activities unrelated to the project. The second criterion was proficiency in the local language.

by field-trainers operating in the *same region* working with the same tribal group — at least in the central and western regions — also lends support to the exogeneity assumption.

Finally, associations that were encouraged to use direct elections have broadly similar *pre-treatment* covariates as those encouraged to use appointments (online appendix, Table A3). Note, however, that APEP did not collect systematic baseline data on the members of the farmer groups. Since over the years some members have left and new members have joined, the current membership body, whose attributes are captured in our representative sample, is different from the one at the time of group formation. For this reason, I only use DC-level pre-treatment variables in regression models that have controls and in the matching analysis I conduct for robustness (see online appendix). A comparison of members' characteristics also suggests that groups receiving different recommendations are comparable.¹⁶ Notwithstanding these findings, to address possible violations of the exogeneity assumption, I conduct a sensitivity analysis for the randomization inference. Discussed formally in the online appendix, this analysis shows that the significance of the study's main results holds even for relatively large violations of the exogeneity assumption.

A major concern regarding the instrument's validity is possible violation of the exclusion restriction. For an unbiased IV estimation, the instrument must affect the outcome only via its impact on the endogenous treatment variable. The concern is that some field-trainers took additional actions that affected the outcomes of interest, and that these actions *are not orthogonal to their leader selection recommendation*. For example, it is possible that trainers who recommended the use of direct elections were more sensitive to accountability concerns and stressed governance issues more forcefully during the facilitation process. There are some reasons, however, to assume that the exclusion restriction holds.

First, there is no evidence to suggest that field-trainers recommending the use of direct elections set up different governance institutions, apart from the manager selection method (online appendix, Table A4). More so, when regressing a summary index of other governance rules on trainer indicators, none of the returned coefficients is significant. This is an important finding

¹⁶ The characteristics of the membership body are, in effect, post-treatment variables that one should not "control" for. A table comparing members' characteristics is not reported here, but can be obtained from the author upon request.

since it suggests that the recommendation to use a leader selection rule is uncorrelated with other governance suggestions, *at the facilitator level*. Finally, due to the large area of coverage and the large number of associations under the purview of each facilitator, the interaction between trainers and groups was very limited once the groups were formed.

To be clear, the identification strategy of this paper is hardly perfect. The small number of trainers and the inability to decouple trainers' attributes from their recommendation or from district effects are obvious limitations. Nonetheless, given the problems associated with asking communities to use random assignment to determine how they are to select their leaders, researchers ought to invest efforts in identifying natural conditions that result in exogenous variation in governance rules. I believe that the instrument used in this study satisfies requirements for IV estimation necessary to justify the analysis presented in the following sections.

6 Measurement of Key Outcome Variables

This study examines the causal effect of direct elections (compared to appointments) on two broad outcomes of interest: (a) *leader responsiveness*, and (b) *members' cooperation*. I operationalize these outcomes by grouping a number of related measures into summary indices. In addition, I use summary indices to measure some of the causal mechanisms that likely mediate the relationship between leader selection rules and leader accountability: (c) *group monitoring*; and (d) *leader network centrality*. Following Anderson (2008), a summary index is a weighted mean of several standardized outcomes, where the weights — the inverse of the covariance matrix — are used to maximize the amount of information captured by the index. This approach improves statistical power while being robust to over-testing because each index represents a single test.

Summary indices ensure that the probability of a false rejection does not increase as additional outcomes are added to the index. More so, summary indices minimize the risk that researchers cherry-pick proxy measures and the risk that researchers misinterpret the importance of individual proxy measures, which may be statistically significant due simply to random chance. Similar to Casey *et al.* (2012), I report the effect of direct elections on both the summary index and the constituent variables that make up the index, but conclude that the treatment (i.e., direct elections) has a causal

effect if and only if the coefficient on the summary index of a “family” of outcomes is statistically significant by conventional standards. When constructing the outcome measures I treat the farmer association as the unit of analysis since this was the level of treatment assignment, taking group averages as necessary. All variables that are analyzed as outcome measures are standardized in units of the control groups; that is, associations assigned to use appointments. The two key outcomes “families” are described in turn.

6.1 *Leader Responsiveness*

Leader responsiveness is defined here as an incumbent’s behavior that is consistent with the preferences of community members (Przeworski *et al.*, 1999). To construct the summary index of leader responsiveness I use four variables capturing actions that group leaders may take, which are valued by community members in the study’s context. The first variable indicates whether members report that the DC manager provides them receipts when selling their crops via the farmer group. In low-income countries such as Uganda, credit-constrained farmer groups are unable to pay cash-on-delivery. Instead, farmers receive payments about four to eight weeks after delivering their crops to their group. To reduce the ability of the DC manager to exploit his position, members have a strong preference to receive a receipt specifying the date and volume of such transactions. Receipts are, therefore, a good proxy of leaders’ responsiveness and accountability.

The second and third measures of responsiveness capture the extent to which the DC manager coordinates sanctioning efforts to enforce cooperative behavior. Recall that the main responsibility of the DC manager is to organize the collective marketing of members’ produce. Collective marketing is, however, subjected to collective action problems. Most importantly, members may refrain from selling their crops through their farmer group. Prior to an establishment of a farmer group, smallholder producers are restricted to selling their crops through middlemen, who can exploit asymmetries in information and bargaining power to offer farmers below-market prices. By organizing, farmers usually obtain higher prices by increasing their bargaining power and by reducing buyers’ transaction costs. Once a farmer group is in place, however, middlemen tend to raise their prices to remain competitive. Importantly, members have a private interest in selling to middlemen, who unlike most farmer groups collect crops at the farmers’ gate, and rarely check for quality. The private gain of selling to middlemen (“defecting”),

however, is conditional on a sufficient number of *other members* selling their crops via the farmer group (“cooperating”). This is because the price offered by middlemen depends on the price that the farmer group secures, which itself crucially depends on volume. If too many members defect, collective marketing collapses altogether.¹⁷ Thus, DCs forbid members from selling to middlemen, and it is the DC manager’s responsibility to enforce this rule by coordinating the sanctioning of defectors.

More so, community members may have a private interest to refrain from making quality-enhancing investments, if *other members* are making such investments, so long as they can “mix” their crops with that of others as part of collective marketing. Since the price that the manager can secure from buyers depends not only on *volume* but also on *quality*, farmer associations commonly pass resolutions requiring members to use best-agricultural practices. The DC manager is responsible to enforce those rules and bylaws. The variables *Member warned: side-selling* and *Member warned: bad agricultural practices* are self-reported indicators derived from the members’ survey.

The fourth measure of leader responsiveness is a binary variable capturing whether members report that the DC is “very transparent.” The variables that make up the leader responsiveness summary index are positively correlated with a Cronbach’s alpha of 0.77.

6.2 Member Cooperation

The cooperation summary index consists of seven variables, out of which four measure the willingness of members to contribute funds to support the capitalization of their community organization. The ability to raise revenues from members is key for the development of community organizations, as it allows increasing the quality and range of services they offer. The first variable is binary measuring whether members express willingness to increase the commission paid to the DC for organizing collective marketing. The second and third variables are self-reported binary measures of whether members paid a fee when joining the farmer group and whether they paid annual dues in the past 12 months. The fourth variable *commitment experiment*, is continuous, capturing the share of members’ payoff from participating in the data collection activities, which they agreed to contribute to their farmer

¹⁷ In a companion paper, I show that farmer associations that manage to overcome this social dilemma have a strong, positive and significant impact on their members’ welfare (Grossman and Hanlon, forthcoming).

association when the research team offered to match their contribution at the end of the enumeration day.

The cooperation index consists also of three behavioral measures, derived from the members' survey, capturing various agriculture practices that are valued by group members. The fifth variable that makes up the cooperation index is the share of a member's seasonal coffee yield sold via their farmer group. The sixth variable is binary indicating whether members dry their coffee beans on tarps rather than on the ground, which represents a costly investment in quality enhancing agricultural "best practices." Recall that collective marketing and investments in quality are subjected to collective action problems and hence proxy well members' cooperation. Lastly, the cooperation index includes an indicator measuring whether members report planting coffee seedlings in the past year. Since it takes 4–5 years until a coffee seedling bears marketable beans, the planting of seedlings demonstrates commitment to growing coffee in the foreseeable future. The variables that make up the cooperation index are positively correlated with a Cronbach's alpha of 0.71.

6.3 Estimation

I use OLS to estimate an Intent-to-Treat (ITT) effect. The ITT model considers Z_j , an exogenous binary variable denoting a field-trainer recommendation to group j , such that $Z_j = 0$ denotes that community organization j was encouraged to use an appointment rule, and $Z_j = 1$ denotes that it was encouraged to use direct elections. Specifically

$$y_j = \alpha_y + \gamma Z_j + \varepsilon_j,$$

where y_j is a standardized outcome variable, γ is the ITT estimate, and ε_j is the error term. The advantage of such analysis is that farmer associations are comparable, if one accepts the encouragement study design. The disadvantage is that it estimates the effect of encouraging farmer associations to use direct elections, not the effect of direct elections itself. To calculate the latter, I use a two-step, generalized method of moments, instrumental variable estimator.

The IV model considers the effect of the leader selection rule — endogenously chosen binary treatment d_j , $E[\varepsilon|d \neq 0]$ — on outcome y_j , conditional on the instrument Z_j ; the field-trainer's recommendation. Note that $d_j = 0$ denotes that the farmer association uses, *in effect*, an appointment rule to

select its leader, while $d_j = 1$ denotes the actual use of direct elections. In some specifications I also control for X_j , denoting a matrix of exogenous covariates.¹⁸ The primary interest is in the regression function:

$$y_j = \alpha_y + \delta d_j + \beta X_j' + \varepsilon_j,$$

where δ is the LATE estimate, β is a vector of coefficients, and ε_j is the error term. The binary decision to obtain the treatment d_j is modeled as the outcome of a linear function of the instrument Z_j and a random term. To compute p -values I use randomization inference for which no assumptions regarding the distributions or independence of potential outcomes are needed.¹⁹

7 Main Results

I turn to report the study's main findings. I begin by testing the study's central hypothesis that direct election increases leader responsiveness. I then turn to test the impact of direct elections on community members' cooperative behavior. ITT and IV results are reported in Table A1 with and without DC-level controls.²⁰ Non-parametric matched pairs ITT and IV estimations, reported in the online appendix (Table 6), yield similar results. In Figures 1 and 2, I provide graphical representation of the ITT estimates with no controls, including 90% confidence intervals.

7.1 Leader Responsiveness

As Figure 1 makes clear, I find that direct elections have a large positive and significant effect on leader responsiveness. The ITT estimate of direct elections on the responsiveness summary index is 0.66 standard deviations

¹⁸ Group controls include years since APEP began operating in the parish, the age of the association, and the number of members when the association was originally founded.

¹⁹ The hierarchical nature of the data, together with the relatively small number of groups ($n = 50$) present some estimation challenges. Using a series of Monte Carlo simulations, (Rader, 2012) demonstrates that when the number of groups is relatively small ($n \leq 50$), randomization inference outperforms cluster-robust standard errors estimation, as the asymptotic properties of the latter are relevant only when the number of groups is large.

²⁰ I report ITT regressions results with Heteroscedastic robust standard errors. Fitting OLS with additional clustering at the district level or the field-trainer produce *smaller* standard errors and are thus less conservative. Note that the ITT estimates are about 0.73 of the effect of the IV estimates, that is, the compliance rate.

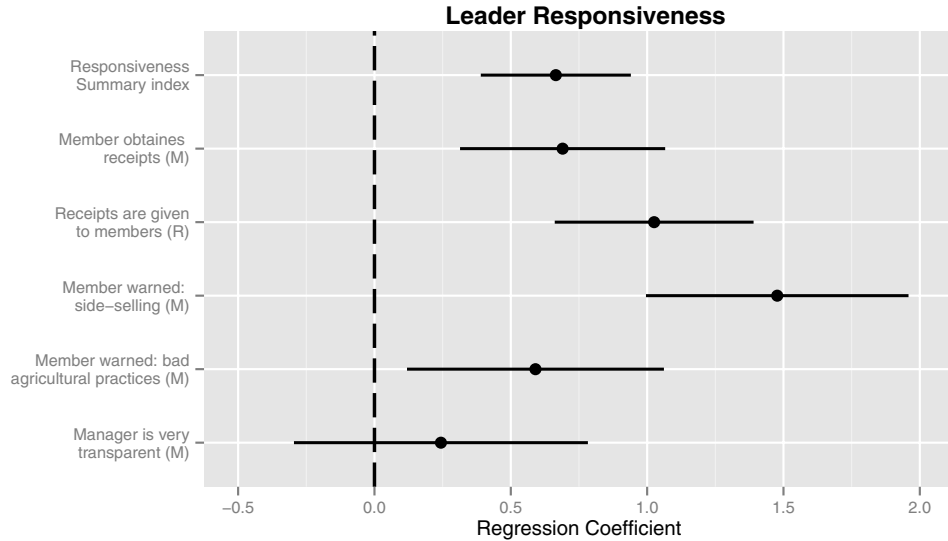


Figure 1. Leader Responsiveness. ITT effects of moving from appointments to direct elections for outcome variables measuring the responsiveness of DC leaders. N . observations = 50.

(p -value = 0.000). Turning to the variables that make up the summary index, members in associations using direct elections are more likely to reports obtaining receipts from the DC manager when selling in bulk. I corroborate this finding with results that are derived from the council representatives' survey (R), which produce similar results. Group members are also significantly more likely to report being sanctioned for side-selling and for not using agricultural "best-practices." Finally, members in direct elections associations are more likely to claim that their manager is transparent, though this difference is not significant (randomization inference p -value = 0.429).

My finding that direct elections result in more responsive office holders is consistent with studies of U.S. state-wide public-officials. For example, consumer-price data suggest that elected regulators produce more pro-consumer policies than appointed regulators, in areas as diverse as public utilities, telecommunications and insurance (Besley and Coate, 2003). Similarly, Huber and Gordon (2004) find that elected judges become more punitive the closer they are to standing for reelection, independent of personal and ideological attributes. They interpret this as a sign that elections increase judges' responsiveness to the public.

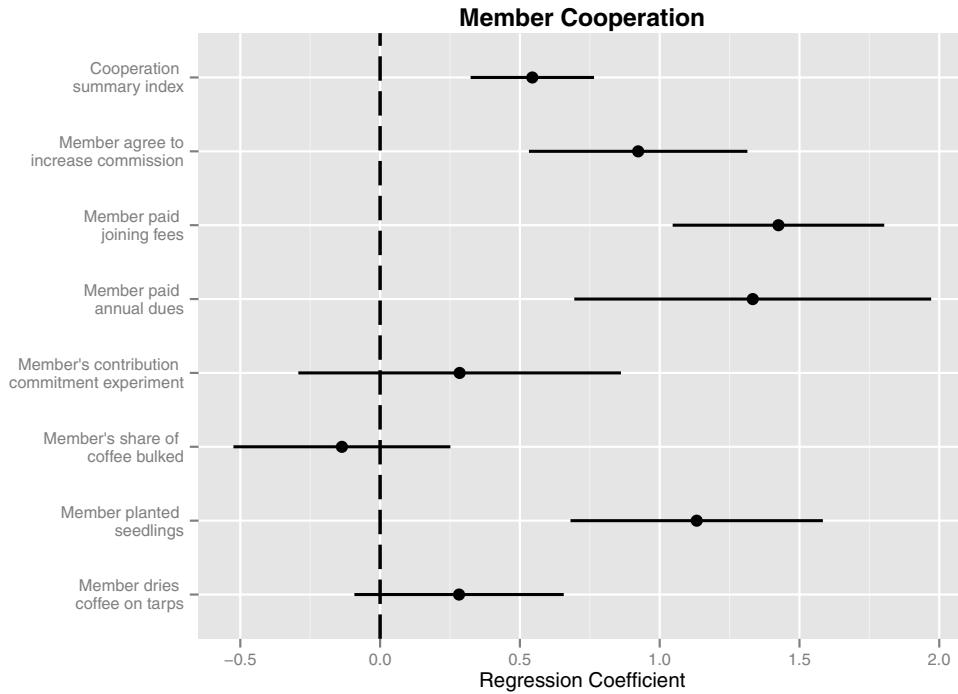


Figure 2. Members' Cooperation. ITT effects of direct elections on outcome variables measuring the cooperative behavior of community members. N . observations = 50.

7.2 Cooperation

Do leader selection rules impact the cooperative behavior of community members? As Figure 2 and Appendix, Table A1 (second panel) show, I find that direct elections have a large, positive and significant effect on members' cooperation; the ITT estimate of direct elections on the cooperation summary index is 0.54 standard deviations with p -value = 0.000. Specifically, I find that community organizations using direct elections are more likely to charge joining fees from new members and to collect annual fees from existing members. In addition, members that vote directly for their DC manager are more likely to agree to pay their group a higher service commission for organizing collective marketing.

Importantly, I find strong evidence that the increase in revenues from members' contributions translate into better services; groups using direct

elections are able to secure higher coffee prices, and are significantly more likely to offer loans, to organize input procurement and pre-season crop purchase (Appendix, Table A1, third panel). Notably associations using direct elections also report a significant increase in membership growth-rate since foundation.

The study’s findings — that the more leaders are accountable and responsive, the more citizens are willing to contribute towards collective goals — are consistent with cross-national studies that have shown that citizens that are allowed to elect leaders in free and fair elections are more willing to pay taxes (Li, 2006), which then finance a higher level of public goods provision (Stasavage, 2005). Our findings are also consistent with results from public goods experiments in the United States (Hamman *et al.*, 2011), Uganda (Grossman and Baldassarri, 2012), and Liberia (Fearon *et al.*, 2009).

8 Mechanisms

Direct elections associations result in more responsive and accountable leaders, members that are more willing to contribute to public goods, and higher levels of goods and services. In this section, I turn to analyze the explanatory power of several mechanisms, described above, that likely mediate the relationship between selection rules and leader responsiveness. First, I test whether elected leaders are more constrained in their behavior because direct elections associations result in higher levels of monitoring. Here special attention is given to separating between horizontal and vertical accountability mechanisms (*H1* and *H2*). I then test whether endowing community members with the power to elect their manager directly results in different “type” of local leaders. Here I distinguish between tangible attributes such as education, cognitive skills, and wealth (*H3*) and the more abstract attribute of commitment to the welfare of community members (*H4*).

8.1 Auditing Institutions and Monitoring Practices

I use nine variables to calculate a summary index of a farmer organization’s level of monitoring. The first measure, derived from the survey of the executive committee members, captures the number of council meetings attended by each council representative in the past 12 months. The second is a three category variable measuring the amount of influence council representatives feel they have over decisions taken at the association level. The third is a

three category variable, derived from the representatives' survey, measuring the extent to which the farmer association uses external accountants to audit its books (*external auditing*). The fourth is a self-reported binary measure of whether council representatives report ever requested to personally review their association's books and records (*internal auditing*). The fifth and sixth variables are self-reported binary measures of whether representatives and ordinary members attended the last general assembly. These measures are closely related to monitoring since the general assembly meeting is devoted mainly for reporting past outcomes and for deliberating future plans.

The monitoring index includes, in addition, three perceptual variables measuring the extent to which members and council representatives view the DC manager as constrained. The seventh and eighth variables derived from the representatives and members' surveys respectively, are binary capturing whether respondents believe that the manager's actions are appropriately monitored. The ninth variable captures whether members hold that the DC manager is accountable. The variables that make up the monitoring index were averaged to the group level. They are all positively correlated with a Cronbach's alpha of 0.82.

As Figure 3 shows, in the Ugandan farmer associations studied herein, the more closely the behavior of a leader is monitored the more likely he is to be responsive to community members.

I now move to examine more closely the relationship between leader selection rules and an organization's levels of monitoring. Consistent with my theoretical expectation, I find a positive and significant impact of direct elections on the farmer association's monitoring level (Appendix, Table A2). The ITT estimate of moving from appointments to direct elections on the monitoring summary index is 0.622 standard deviations (p -value = 0.002).

I have argued that a high level of group monitoring can be the results of either (a) vertical accountability mechanism in the form of enhanced grassroots participation ($H1$) and/or (b) horizontal accountability mechanism in the form of lax oversight by council representatives ($H2$). Importantly, the evidence at hand strongly suggests that the positive impact of direct elections on monitoring is more likely due to lower investment in monitoring by *council representatives* in appointments associations than to greater grassroots monitoring by *ordinary members* in direct elections associations.

I do not find any evidence that direct elections lead to greater grassroots participation. Using self-report measures, I find that members in direct

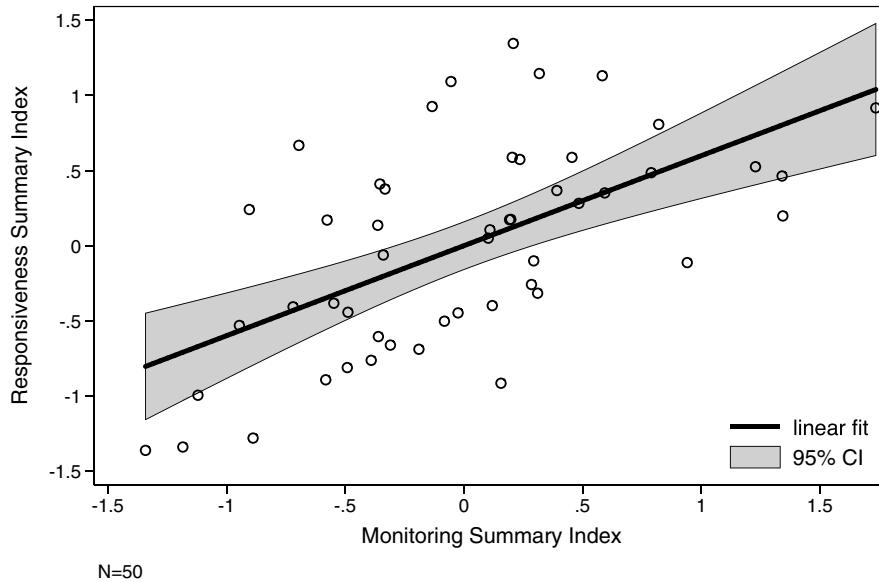


Figure 3. Relation between monitoring and leader responsiveness.

election associations are no more likely to attend the last general assembly meeting than their counterparts in associations using appointments (p -value = 0.218). Importantly, they are also no more likely to participate in training sessions (p -value = 0.236) nor in group meetings at the village level (p -value = 0.590).²¹ These findings are notable for two main reasons. First, they point to the fact that community members are reluctant to invest in costly monitoring, even when given the opportunity to vote for their preferred candidate. Secondly, they help to alleviate concerns that APEP field-trainers recommended the use of direct elections in areas that are already somehow more “participatory.”

Indeed, the study’s findings are more in line with a *horizontal accountability* mechanism than a *vertical* one. The variables that make up the monitoring summary index, which are associated with the overseeing role of council representatives, tend to exhibit large treatment effects. Compared to associations using appointments, associations using direct elections are more likely

²¹ These findings are robust to whether using the *number* of meetings/training sessions a member attended in the past season, the *share* of meetings or training sessions in the past year, or the *last meeting or training session*.

to hire licensed accountants to audit the DC's books (p -value = 0.004). Council representatives in associations using direct elections are also more likely to report having meaningful influence over decisions made at the association level (p -value = 0.029), to attend council meetings (p -value = 0.129) and the last general assembly meeting (p -value = 0.002). Importantly, they are also more likely to request to inspect the DC's books (internal auditing; p -value = 0.159).

Though interesting, the above findings nonetheless leave open the theoretical question of *why* direct elections associations result in greater horizontal accountability. Indeed, more in-depth analysis is needed to tease out the reasons behind the observed difference in the behavior of council representatives across leader selection rules. I now turn to directly test for the presence of a '*substitution*' effect: the idea that direct elections associations result in higher levels of monitoring because monitoring institutions and 'closeness' between appointed and appointers are seen as substitutes.

8.2 Network Centrality and Substitution Effects

To test the "substitution" hypothesis, I collected complete network data on the friendship ties among council representatives in each of the 50 sampled farmer associations.²² Using a simple roster, each council representative was asked whether he or she was a close friend of *each of the other council members*.²³ If a substitution effect is to explain the variation in monitoring levels between direct elections and appointments, the following should be observed: (i) leaders of farmer associations using appointments have a more central position in the friendship network, and (ii) farmer associations in which leaders have a more central position in the friendship network exhibit lower levels of monitoring.

I use four variables to construct a summary index of the centrality of the DC manager within the council friendship network. One natural proxy measure is *degree centrality*: the number of direct connections a node has. Degree centrality builds on the basic intuition that a prominent network actor is one that has high involvement in many relations (Freeman, 1979).

²² The term "complete" refers to the fact that the council is a clearly defined body with transparent membership rules, and data was collected on all network nodes.

²³ Following a sorting question, respondents were asked: "Is [NAME] a close friend or do you just know him or her? by close friend, I mean that you (a) eat together regularly; (b) you can leave your child with him or her if you need to travel for several days; and (c) he or she will help you in case of a family death."

Since friendship ties are directional, I calculate separately both indegree and outdegree centrality measures. Indegree centrality (or “prestige”) captures the extent to which an actor in a network “serves as the object” of relations sent by other actors (normalized by the number of direct ties).²⁴ Outdegree centrality captures, instead, the number of endpoints adjacent to a node. The third proxy-measure of centrality is a node’s *closeness*, which builds on the intuition that an actor is more central the lower his or her total distance to all other network actors (Wasserman and Faust, 1994).²⁵ Finally, I calculate *Eigenvector centrality*, a centrality measure that builds on the idea that ties to high-scoring nodes should contribute more to the centrality score of a node than equal ties to low-scoring nodes.

Consistent with the substitution hypothesis, I find strong evidence suggesting that appointed leaders are more central to the friendship network than directly elected leaders (Appendix, Table A3, second panel). The ITT estimate of direct elections on the leaders’ friendship centrality summary index is -0.671 standard deviations (p -value = 0.006). To demonstrate a substitution effect, it must also be shown that the more central to the elite friendship network a leader is, the less likely he is to be monitored. As Figure 4 shows, a negative correlation is observed in the Ugandan farmer associations studied herein.

One could argue that close friendship ties between community elites can substitute for costly formal monitoring institutions without adverse effects. The data, however, suggests otherwise. On one hand, council representatives in associations using appointments do not view the DC manager as less monitored or accountable than council representatives in direct elections associations (p -value = 0.409). On the other hand, as shown in Table A2, ordinary members in association using appointments view the leader as significantly less accountable (p -value = 0.001) and less monitored (p -value = 0.023) than their counterparts in direct elections associations. In other words, the perception of members and representatives in appointment associations with respect to the accountability of the group leader diverges, and this divergence is consequential. In sum, the paper’s findings are consistent with a model in which the close friendship ties between appointed officials

²⁴ More formally, indegree centrality, is defined as $C_D(n_i) = \frac{d_i(n_i)}{g-1}$, where $d_i(n_i)$ is the indegree of a node n_i and g is the number of all nodes in the network.

²⁵ The distance between pairs of nodes is defined by the length of their shortest paths. *Farness* is defined as the sum of a node’s distances to all other nodes, and its “closeness” is defined as the inverse of the farness.

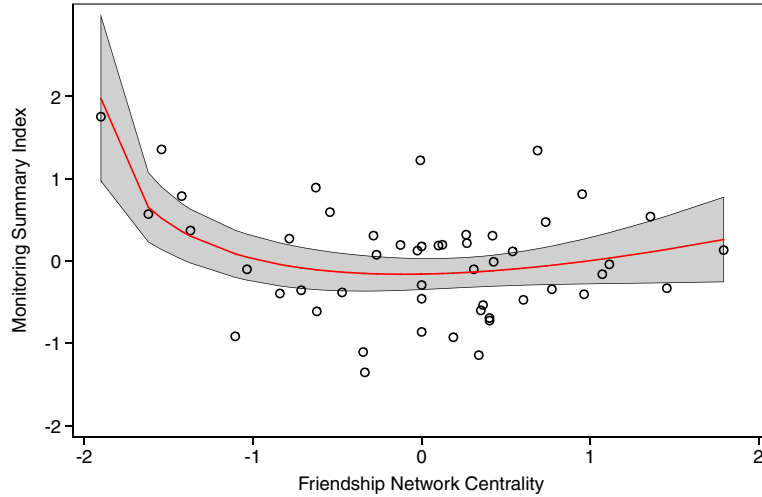


Figure 4. Figure describes the relation between a leader’s friendship network centrality and the associations’ monitoring-level. Fitted line and 95% confidence intervals derived from fractional polynomial regression with bandwidth of 0.5.

and appointers substitute for formal monitoring institutions, leading to loss of confidence by ordinary members and subsequently to a decline in public goods contributions.

8.3 Leader’s Type

I discussed above the relevance of the theory of “bundling” developed by Besley and Coate (2003). If appointment systems are constrained by a bundling effect, we may expect to see that direct elections result in leaders with different attributes than appointed leaders (*H3*). I do not, however, find differences between elected and appointed leaders in any observable attributes, such as age, land size, wealth, education, or cognitive abilities.²⁶ Results are presented in Table A3, top panel.

Moving to less tangible attributes, an empirical challenge arises when testing whether direct elections result in leaders who are more committed to the

²⁶ The cognitive tests included solving a simple maze in less than two minutes and solving a raven test comprised of 12 questions in two minutes. The wealth measure was constructed using questions about ownership of 12 assets, such as bicycles and livestock, which reflect the purchasing power of farmers in rural Uganda.

welfare of community members (*H4*). This is because to avoid circularity, I must demonstrate that elected leaders are somehow more committed to members' welfare independent of their level of responsiveness. To overcome this challenge, I employ a series of behavioral experiments.

The first experiment is a dictator game (DG). In DGs, two anonymous players are allotted a certain endowment — in our case, 10 coins of 100 Ugandan Shillings, equivalent to half a day's wage in rural Uganda. The first player (the “decider”) offers a portion of this sum to a second player who cannot reject the offer (the “recipient”). In this one-shot game, a purely self-interested “decider” would offer zero. It is common to interpret offers in DGs as measures of altruism, as they are not directly linked to kinship, reciprocity, reputation or the immediate threat of punishment (Camerer, 2003). In my version of the DG, DC managers made two allocation decisions: (i) to an anonymous sub-county resident (i.e., “stranger” condition), and (ii) to an anonymous farmer association member (i.e., ‘member’ condition).²⁷

The results of this experiment are shown in Figure 5 and Table A2, bottom panel. Though directly elected and appointed leaders allocate similar amounts in the “stranger” condition, directly elected leaders allocate to anonymous group members a significantly larger share of their endowment (556 compared to 349 Ugandan Shillings, *p-value* = 0.030). This finding suggests that though both types of leaders have broadly similar other-regarding preferences, directly elected leaders are somehow more concerned about the welfare of group members.

Results from a series of Third Party Punishment experiments (TPPs) are broadly consistent with the DG findings. In TPPs, two players are allotted a sum of real money (10 coins of 100 Ugandan Shillings), and a third player gets one-half of this amount. Player 1 (the “decider”) must decide how much of the stake to transfer to player 2 (a “recipient”) who makes no decisions in the game. Then, *before* learning the actual amount player 1 transferred to player 2, player 3 (“monitor”) must decide whether to pay 1 coin (i.e., 20% of her endowment) to punish player 1, causing player 1 to suffer a deduction of 3 coins, for all possible transfers player 1 might make.²⁸ Since the experiment is an anonymous one-shot interaction, a purely self-interested player 3 would

²⁷ DC managers were told before making their allocation decision that community members are told that they were receiving an allocation from “one of the leaders of the farmer association.”

²⁸ In technical terms, I employ the strategy set protocol (Henrich *et al.*, 2004), in which the punishment strategy of player 3 is elicited *ex-ante* for all possible offers.

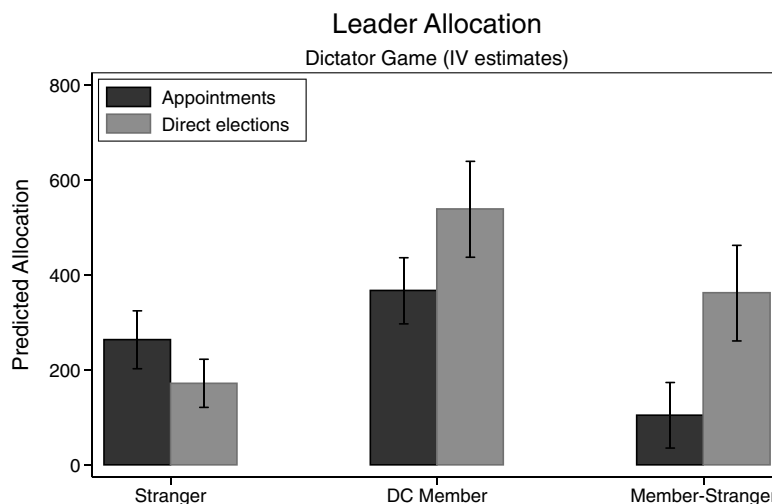


Figure 5. Figure describes the predicted allocations of managers in a dictator game in Ugandan Shillings to (i) strangers and (ii) DC members, as well as (iii) the predicted difference between those two allocations.

not pay to punish player 1. Knowing this, a purely self-interested player 1 would offer zero to player 2. Since the payoff of player 3 depends only on her punishment decisions, sanctioning provides a direct behavioral measure of the willingness to suffer a personal cost in order to enforce a norm of fairness.

In my version of the TPP, the decider and the passive recipient were ordinary group members randomly selected to their roles, and the “monitor” was one of the four members of the DC executive committee.²⁹ I find that in all unequal allocations, directly elected leaders are more likely to suffer a personal cost to punish “selfish” members violating a within-group norm of cooperation, though not reaching significance level (Appendix, Table A2). These results are consistent with the survey findings reported above: that members in direct elections associations are more likely to report being sanctioned for producing low quality coffee and for side-selling to middlemen.

That directly elected leaders, in both experiments, were more willing to suffer monetary costs under conditions of anonymity, suggests that the

²⁹ Executives knew that players 1 and 2 are members, members knew that player 3 was “one of the DC leaders,” and player 3 knew what players 1 and 2 knew.

greater concern expressed by elected leaders to the welfare of community members *cannot be easily reduced to reelection considerations*. The DG and TPP do not tell us, however, whether leaders were elected because they were already more committed to members' welfare or whether they are reciprocating for being elected. Nonetheless, the fact that elected and appointed leaders contributed similarly to "strangers" suggest reciprocal relations between leaders and "followers" that are triggered by the electoral process itself.

9 Discussion and Conclusion

This paper studies the impact of leader selection rules in the context of community associations. Past studies on the impact of formal governance institutions are commonly constrained by pervasive identification problems. I address this challenge by taking advantage of unique natural conditions that resulted in plausibly exogenous variation in the rules for selecting leaders of Ugandan farmer associations. Results presented here suggest that direct elections have a positive, substantial, and significant effect on the extent to which local leaders are responsive to community members and on the willingness of community members to contribute toward local public goods. There is increased confidence in these findings because the results are consistent across a large number of behavioral and perceptual measures and a variety of different data sources. This study demonstrates that contrary to the prevailing view that existing hierarchical structures are resilient, exogenously introduced governance institutions matter, even in intimate and information-rich local environments. This finding has important theoretical implications.

From a theoretical perspective the study's main finding sheds light on the appropriate way to model political accountability at the community level in low-income countries.³⁰ The use of non-cooperative games, which pay close attention to formal rules and procedures, has been the common framework to model political accountability at the national level (e.g., Fearon (1999)). This framework, however, is of little use if local elites can always work around formal rules, for example by taking advantage of their ability to form coalitions and make credible promises and threats. If informal rules are resilient, then cooperative game theory may be a more appropriate

³⁰ I thank Macartan Humphreys for raising this point.

framework. As is well known, cooperative games focus on the outcomes that result when actors come together in different combinations, allowing the possibilities for binding agreements, meaningful bargaining and communication. Whereas the results of a non-cooperative analysis largely depend on the precise form of the formal rules, cooperative games assume that negotiations are amorphous, and that formal rules are of little use since it is hard to know which *formal and informal* rules are at all relevant. In this study, I find that both community leaders and community members are responding to changes in formal rules in predictable ways, demonstrating the relevance of non-cooperative political accountability models even in extremely localized contexts.

Turning to the factors that mediate the relationship between leader selection rules and leader accountability, I find that in community organizations using appointments, appointers tend to select candidates with whom they interact frequently and whom they view as friends. In such cases, the closeness between appointed officials and appointers (here, council representatives) provide a disincentive for the latter to invest in costly monitoring. I refer to this process as a *substitution effect*. Faced with low levels of monitoring, appointed office holders, who occupy a central position in the local elite network, are less constrained and, therefore, less responsive to community members. Consequently, community members are less likely to contribute toward collective goals, resulting in lower levels of group services. In addition, I find that direct elections result in leaders who are more likely to consider the welfare of members, and therefore to be more responsive to the signaled preferences of community members.

The mechanisms identified in this paper have important implications to our understanding of how the interaction between informal ties and formal institutions impacts political accountability in community settings. Specifically, my findings complement and expand the pioneering work by Tsai (2007) that shows that village-wide leaders who are “embedded” in encompassing solidarity groups are *more* accountable to community members. Conversely, I find that leaders who occupy central positions within the network of village elite are *less* likely to be monitored, resulting in inferior public goods outcomes. Whereas the focus of Tsai (2007) is on *vertical* relations — in her study, village leaders and members are implicitly theorized as part of the same social network — this study highlights, instead, the importance of *horizontal* relations among village elites.

The study's findings that *bottom up grassroots monitoring* efforts are of limited effectiveness, which is consistent with Olken (2007), has also important policy and a theoretical implications. In response to the findings reported in Bjorkman and Svensson (2009), development practitioners are currently making considerable investments in facilitating and strengthening grassroots monitoring institutions. The grassroots vertical monitoring approach, however, underplays free-riding problems, while assuming that community members are willing to both invest time and efforts in costly monitoring and confront their community leaders. Indeed, my finding that community members are reluctant to monitor community leaders, even in direct elections associations, questions the sustainability of the grassroots vertical monitoring approach in the medium to long-term.

The study's key findings contribute also to the ensuing debate regarding the viability of governance reforms that seek to empower citizens vis-à-vis local leaders, on which the entire CDD model is premised. On one hand, the main findings presented here are consistent with results from studies of formal governance institutions in Liberia (Baldwin and Mvukiyehe, 2011), Mexico (Díaz-Cayeros *et al.*, 2012) and China (Luo *et al.*, 2007). On the other hand, the paper's findings are in tension with recent studies by Casey *et al.* (2012) and by Humphreys *et al.* (2013) that find negligible effects of exogenously induced governance institutions in the context of CDD projects in Sierra Leone and the DRC.

Reconciling those findings is an important avenue for future research. Here I can only offer preliminary thoughts regarding the *type of institutions* and the *characteristics of community organizations* that can increase the likelihood of a positive impact. First, there is some evidence that procedural, technical rules, when followed, are likely to be consequential. For example, notwithstanding the mixed evidence regarding the effectiveness of the CDD model, studies in Afghanistan (Beath *et al.*, 2013a) and Indonesia (Olken, 2010) demonstrate the importance of the procedural rules used for selecting development projects at the community level. Both studies randomly assign whether communities use village meetings convened by village councils to select development projects or a secret-ballot referenda that give villagers the opportunity to directly vote for their preferred project. The finding of both studies, that procedural rules impact the extent to which elite are able to control resource allocation, is consistent with the findings of this paper, as well as with work on gender quotas in India (Chattopadhyay and Duflo, 2004).

Secondly, the type of community organization may also be consequential for the impact of formal governance institutions. For example, there are good reasons to assume that formal governance institutions are more likely to have a bite in *voluntary groups*. First, members of voluntary groups, such as farmer, credit and women associations, usually have credible exit options, whereas villagers do not. Exit options tends to magnify voice (Hirschman, 1970). Secondly, in voluntary organizations the preferences of members with respect to the goods produced by the leader are closely aligned. Thus voluntary community organizations tend to be more internally cohesive than the broader villages in which they are embedded, where participatory processes are likely to interact with in-group tensions (Mansuri and Rao, 2013).

It seems plausible that the composition of the membership body is also relevant to this discussion. Farmer associations, for example, are designed to secure higher output prices and lower input prices, and are *not* designed to necessarily address the needs of marginalized groups. In fact, due to high membership fees and demands on members' time, women and the very poor tend to be under represented in the membership body. This means that members of farmer associations, generally men with access to land, have more resources in their disposal to affect desirable change.

Finally, the findings of this paper raise important concerns regarding generalizability to the wider institutions literature, which is interested in the broad question of how democratic institutions affect growth, development, or accountability. Much of the empirical research on institutions has used historical and qualitative methods (e.g., Putnam, 1993) that suffer from identification problems, or instrumental variables approaches using cross-national data (e.g., Acemoglu *et al.*, 2001) that also suffers from a myriad of empirical challenges (Pande and Udry, 2006). These well-known limitations have sparked work on the consequences of institutions at the community level, using field and natural experiments methods.

But though micro-level studies have strong internal validity, the generalizability of their findings is an open question. Thus micro-level researchers should be more specific about what aspects of the institutions under study illuminates micro-foundational, or universal, behavior and what part is context specific. On the other hand, there are good reasons to believe that the two mechanisms identified in this paper — that appointment systems are subjected to substitution effects and that direct election, by overcoming bundling effects, result in leaders that are more committed to members' welfare — capture micro-foundational behaviors. On the other hand, a key

challenge in generalizing from community-level studies is the fact that the relationship between community leaders and members is multi-dimensional whereas in national politics the relationship between incumbent and constituents is limited to representation and perhaps constituency service. In sum, the extent to which the results of this study and other community-level studies can contribute to the broader institutions literature is still an open question for future research to address.

References

- Acemoglu, D., S. Johnson, and J. A. Robinson. 2001. "The Colonial Origins of Comparative Development: An Empirical Investigation." *The American Economic Review* 91(5): 1369–1401.
- Alesina, A. and E. La Ferrara. 2005. "Ethnic Diversity and Economic Performance." *Journal of Economic Literature* 43(3): 762–800.
- Anderson, M. L. 2008. "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." *Journal of the American Statistical Association* 103(484): 1481–1495.
- Angrist, J. D. and J.-S. Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton: Princeton University Press.
- Araujo, M. C., F. H. G. Ferreira, P. Lanjouw and B. Özler. 2008. "Local Inequality and Project Choice: Theory and evidence from Ecuador." *Journal of Public Economics* 92(5–6): 1022–1046.
- Ashraf, N., X. Gine, and D. Karlan. 2009. "Finding Missing Markets (and a Disturbing Epilogue): Evidence from an Export Crop Adoption and Marketing Intervention in Kenya." *American Journal of Agricultural Economics* 91(4): 973–990.
- Baldassarri, D. and G. Grossman. 2013. "The Effect of Group Attachment and Social Position on Prosocial Behavior. Evidence from Lab-in-the-Field Experiments." *PLoS ONE* 8(3): e58750.
- Baldwin, K. and E. Mvukiyehe. 2011. "The Effects of Participatory Processes for Selecting Leaders: Evidence from Changes to Traditional Institutions in Liberia." Unpublished manuscript.
- Bardhan, P. 2002. "Decentralization of Governance and Development." *The Journal of Economic Perspectives* 16(4): 185–205.
- Bardhan, P. and D. Mookherjee. 2006. "Decentralisation and Accountability in Infrastructure Delivery in Developing Countries." *The Economic Journal* 116(508): 101–127.
- Bates, R. H. 1981. *Markets and States in Tropical Africa: The Political Basis of Agricultural Policies*. Berkeley: University of California Press.
- Beath, A., F. Christia, and R. Enikolopov. 2013a. "Direct Democracy and Resource Allocation: Experimental Evidence from Afghanistan." *MIT Political Science Department Research Paper* 2011-6.
- Beath, A., F. Christia, and R. Enikolopov. 2013b. "Do Elected Councils Improve Governance? Experimental Evidence on Local Institutions in Afghanistan." *MIT Political Science Department Research Paper* 2013–24.
- Bernard, T., M.-Hélène Collion, A. de Janvry, P. Rondot, and E. Sadoulet. 2008. "Do Village Organizations Make a Difference in African Rural Development? A Study for Senegal and Burkina Faso." *World Development* 36(11): 2188–2204.
- Besley, T. and S. Coate. 2003. "Elected Versus Appointed Regulators: Theory and Evidence." *Journal of the European Economic Association* 1(5): 1176–1206.

- Bjorkman, M. and J. Svensson. 2009. "Power to the People: Evidence from a Randomized Field Experiment on Community-Based Monitoring in Ugnada." *Quarterly Journal of Economics* 124(2): 735–769.
- Bowles, S. and H. Gintis. 2004. "Persistent Parochialism: Trust and Exclusion in Ethnic Networks." *Journal of Economic Behavior & Organization* 55(1): 1–23.
- Bueno de Mesquita, B., A. Smith, R. M. Siverson, and J. D. Morrow. 2003. *The Logic of Political Survival*. Cambridge, MA: MIT Press.
- Camerer, C. 2003. *Behavioral Game Theory: Experiments in Strategic Interaction*. New York, NY: Russell Sage Foundation.
- Casey, K., R. Glennerster, and E. Miguel. 2012. "Reshaping Institutions: Evidence on External Aid and Local Collective Action." *Quarterly Journal of Economics* 127(4): 1755–1812.
- Cavazos, R. J. 2003. "The Political Economy of Electricity Deregulation: Appointed vs. Elected Utility Commissioners." *Review of Policy Research* 20(2): 255–262.
- Chattopadhyay, R. and E. Duflo. 2004. "Women as Policy Makers: Evidence from a Randomized Policy Experiment in India." *Econometrica* 72(5): 1409–1443.
- Dasgupta, A. and V. A. Beard. 2007. "Community Driven Development, Collective Action and Elite Capture in Indonesia." *Development and Change* 38(2): 229–249.
- Díaz-Cayeros, A., B. Magaloni, and A. Ruiz Euler. 2012. "Traditional Government, Citizen Engagement, and the Distribution of Public Goods: Local Evidence from Mexico." *Working Paper*.
- Dunning, T. 2008. "Model Specification in Instrumental-Variables Regression." *Political Analysis* 16(3): 290–302.
- Easterly, W. 2006. *The White Man's Burden: Why the West's Efforts to Aid the Rest Have Done So Much Ill and So Little Good*. New York, NY: Penguin Press.
- Fearon, J. D. 1999. Electoral Accountability and the Control of Politicians: Selecting Good Types versus Sanctioning Poor Performance. In *Democracy, Accountability, and Representation*, Adam Przeworski, Susan C. Stokes, and Bernard Manin (eds.), Cambridge: Cambridge University Press, pp. 55–97.
- 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* 99(2): 287–291.
- Freeman, Linton C. 1979. "Centrality in Social Networks Conceptual Clarification." *Social Networks* 1(3): 215–239.
- Fritzen, S. A. 2007. "Can the Design of Community-Driven Development Reduce the Risk of Elite Capture? Evidence from Indonesia." *World Development* 35(8): 1359–1375.
- Grossman, G. and D. Baldassarri. 2012. "The Impact of Elections on Cooperation: Evidence from a Lab-in-the-Field Experiment in Uganda." *American Journal of Political Science* 56(4): 964–985.
- Grossman, G. and W. W. Hanlon. forthcoming. "Do better monitoring institutions increase leadership quality in community organizations? Evidence from Uganda." *American Journal of Political Science*.
- Gugerty, M. K. and M. Kremer. 2008. "Outside Funding and the Dynamics of Participation in Community Associations." *American Journal of Political Science* 52(3): 585–602.
- Hamman, J. R., R. A. Weber, and J. Woon. 2011. "An Experimental Investigation of Electoral Delegation and the Provision of Public Goods." *American Journal of Political Science* 55(4): 738–752.
- Hanssen, F. A. 1999. "The Effect of Judicial Institutions on Uncertainty and the Rate of Litigation: The Election versus Appointment of State Judges." *The Journal of Legal Studies* 28(1): 205–232.
- Henrich, J., R. Boyd, S. Bowles, C. Camerer, E. Fehr, and H. Gintis. 2004. *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies*. Oxford: Oxford University Press.
- Hess, F. M. 2008. "Looking for Leadership: Assessing the Case for Mayoral Control of Urban School Systems." *American Journal of Education* 114(3): 219–245.

- Hirschman, A. O. 1970. *Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States*. Cambridge, MA: Harvard University Press.
- Huber, G. A. and S. C. Gordon. 2004. "Accountability and Coercion: Is Justice Blind when It Runs for Office?" *American Journal of Political Science* 48(2): 247–263.
- Humphreys, M., R. S. de la Sierra, and P. van der Windt. 2013. "Exporting Institutions: A Randomized Experiment in Eastern Congo." *Working Paper*.
- Hussi, P., J. Murphy, O. Linberg, and L. Brenneman. 1993. *The Development of Cooperatives and Other Rural Organisations*. Washington, DC: World Bank: World Bank Technical Paper No. 199.
- Khwaja, A. I. 2009. "Can Good Projects Succeed in Bad Communities?" *Journal of Public Economics* 93(7–8): 899–916.
- Krishna, A. 2002. *Active Social Capital: Tracing the Roots of Development and Democracy*. New York, NY: Columbia University Press.
- Labonne, J. and R. S. Chase. 2009. "Who is at the Wheel When Communities Drive Development? Evidence from the Philippines." *World Development* 37(1): 219–231.
- Lele, U. 1981. "Co-operatives and the poor: A comparative perspective." *World Development* 9(1): 55–72.
- Li, Q. 2006. "Democracy, Autocracy, and Tax Incentives to Foreign Direct Investors: A Cross-National Analysis." *Journal of Politics* 68(1): 62–74.
- Lindberg, S. I. 2006. *Democracy and Elections in Africa*. Baltimore, MD: Johns Hopkins University Press.
- Luo, R., L. Zhang, J. Huang, and S. Rozelle. 2007. "Elections, fiscal reform and public goods provision in rural China." *Journal of Comparative Economics* 35(3): 583–611.
- Mansbridge, J. 2003. "Rethinking Representation." *American Political Science Review* 97(04): 515–528.
- Mansuri, G. and V. Rao. 2013. *Localizing Development: Does Participation Work?* Washington, DC: World Bank.
- Martinez-Bravo, M., G. Padro-i Miquel, N. Qian, and Y. Yao. 2011. "Do Local Elections in Non-Democracies Increase Accountability? Evidence from Rural China." *NBER Working Paper 16948*.
- Miguel, E. and M. K. Gugerty. 2005. "Ethnic Diversity, Social Sanctions, and Public Goods in Kenya." *Journal of Public Economics* 89(11–12): 2325–2368.
- Olken, B. A. 2007. "Monitoring Corruption: Evidence from a Field Experiment in Indonesia." *Journal of Political Economy* 115(2): 200–249.
- Olken, B. A. 2010. "Direct Democracy and Local Public Goods: Evidence from a Field Experiment in Indonesia." *American Political Science Review* 104(02): 243–267.
- Pande, R. and C. Udry. 2006. Institutions and Development: A View from Below. In *Advances in Economics and Econometrics: Theory and Applications*, R. Blundell, Newey W. and T. Persson, (eds.), vol. 2, Cambridge, UK: Cambridge University Press, pp. 349–412.
- Pitkin, H. F. 1967. *The Concept of Representation*. Berkeley, CA: University of California Press.
- Platteau, J.-P. and A. Abraham. 2002. "Participatory Development in the Presence of Endogenous Community Imperfections." *Journal of Development Studies* 39(2): 104–136.
- Ponte, S. 2002. *Farmers and Markets in Tanzania: How Policy Reforms Affect Rural Livelihoods in Africa*. Oxford: Heinemann.
- Przeworski, A., S. C. Stokes, and B. Manin. 1999. *Democracy, Accountability, and Representation*. Cambridge, UK: Cambridge University Press.
- Putnam, R. 1993. *Making Democracy Work*. Princeton, NJ: Princeton University Press.
- Rader, K. T. 2012. "Randomization Tests and Inference with Grouped Data." *Unpublished Manuscript*.
- Seabright, P. 1996. "Accountability and Decentralisation in Government: An Incomplete Contracts Model." *European Economic Review* 40(1): 61–89.

- Shaub, M. K., F. Collins, O. Holzmann, and S. H. Lowensohn. 2005. "Self-Interest vs. Concern for Others." *Strategic Finance* 86(9).
- Stasavage, D. 2005. "Democracy and Education Spending in Africa." *American Journal of Political Science* 49(2): 343–358.
- Tsai, L. L. 2007. "Solidary Groups, Informal Accountability, and Local Public Goods Provision in Rural China." *American Political Science Review* 101(02): 355–372.
- Wasserman, S. and K. Faust. 1994. *Social Network Analysis: Methods and Applications*. Vol. 8, Cambridge, UK: Cambridge University Press.

Appendix

Table A1. Main outcome measures.

	RI	ITT		IV		<i>n</i>
	<i>p</i> -value	nc	c	nc	c	
<i>Leader responsiveness</i>						
Responsiveness summary index	0.000	0.665 (0.164)	0.781 (0.156)	0.917 (0.235)	1.106 (0.239)	50
Member obtained receipts (M)	0.009	0.690 (0.224)	0.856 (0.229)	0.951 (0.308)	1.213 (0.350)	50
Receipts given to members (R)	0.000	1.026 (0.217)	1.178 (0.225)	1.413 (0.330)	1.669 (0.378)	50
Member warned: side-selling (M)	0.000	1.477 (0.287)	1.703 (0.262)	2.035 (0.459)	2.412 (0.475)	50
Member warned: bad agri practices (M)	0.044	0.590 (0.281)	0.788 (0.269)	0.814 (0.379)	1.116 (0.363)	50
Manager is very transparent (M)	0.429	0.244 (0.321)	0.223 (0.292)	0.336 (0.433)	0.316 (0.391)	50
<i>Members' Cooperative Behavior</i>						
Cooperation summary index	0.000	0.544 (0.132)	0.614 (0.124)	0.750 (0.200)	0.869 (0.217)	50
Member agree to increase DC commission	0.002	0.923 (0.233)	0.999 (0.256)	1.272 (0.330)	1.415 (0.385)	50
Member paid joining fees	0.000	1.425 (0.226)	1.467 (0.232)	1.963 (0.380)	2.077 (0.401)	50
Member paid annual dues (past season)	0.000	1.333 (0.380)	1.514 (0.355)	1.837 (0.550)	2.144 (0.560)	50
Member's contribution: commitment experiment	0.389	0.285 (0.344)	0.299 (0.365)	0.392 (0.481)	0.423 (0.511)	50
Member's share of coffee bulked	0.611	-0.137 (0.231)	-0.191 (0.240)	-0.188 (0.316)	-0.270 (0.326)	50
Member planted seedlings (past season)	0.000	1.132 (0.269)	1.131 (0.259)	1.560 (0.398)	1.601 (0.385)	50
Member dried coffee on tarps (past season)	0.284	0.282 (0.223)	0.487 (0.266)	0.389 (0.305)	0.690 (0.335)	50
<i>DC Services</i>						
Standard coffee prices per kg (E)	0.007	0.787 (0.241)	0.926 (0.311)	1.114 (0.391)	1.344 (0.469)	41
Collective input procurement (E)	0.118	0.537 (0.335)	0.658 (0.307)	0.740 (0.434)	0.932 (0.405)	50
Loan services to members (E)	0.138	0.620 (0.381)	0.757 (0.358)	0.854 (0.515)	1.072 (0.498)	50
Pre-season purchase (E)	0.260	0.436 (0.364)	0.351 (0.356)	0.607 (0.488)	0.499 (0.487)	49
Membership growth rate	0.084	0.795 (0.471)	0.667 (0.429)	1.195 (0.721)	1.020 (0.634)	38

Heteroscedastic robust standard errors in parentheses.

RI stands for randomization inference, ITT for Intent-to-Treat analysis, IV for Instrumental Variable
n and nc denote regression models with and without controls, respectively.

M refers to members' survey; R to representatives' survey and E to executives' survey, respectively.

Table A2. Causal mechanisms: Group monitoring and network centrality.

	RI	ITT		IV		<i>n</i>
	<i>p</i> -value	nc	c	nc	c	
<i>Monitoring</i>						
Monitoring summary index	0.002	0.611 (0.198)	0.688 (0.177)	0.841 (0.265)	0.975 (0.263)	50
DC council meeting frequency (E)	0.129	0.676 (0.495)	0.700 (0.482)	0.908 (0.653)	0.959 (0.638)	42
Representatives influence over decisions (R)	0.029	0.577 (0.231)	0.658 (0.260)	0.795 (0.325)	0.931 (0.368)	50
External auditing (R)	0.004	1.460 (0.589)	1.659 (0.530)	2.012 (0.778)	2.350 (0.745)	50
Internal auditing (R)	0.159	0.409 (0.276)	0.430 (0.290)	0.564 (0.378)	0.610 (0.408)	50
General assembly: reps' attendance (R)	0.002	1.057 (0.344)	1.153 (0.336)	1.457 (0.478)	1.633 (0.509)	50
General assembly: members' attendance (M)	0.218	0.404 (0.347)	0.554 (0.329)	0.557 (0.463)	0.784 (0.434)	50
Manager is monitored (R)	0.409	0.307 (0.411)	0.388 (0.332)	0.423 (0.550)	0.550 (0.453)	50
Manager is monitored (M)	0.023	0.723 (0.323)	0.792 (0.301)	0.997 (0.421)	1.122 (0.406)	50
Manager is accountable (M)	0.001	0.975 (0.285)	1.065 (0.286)	1.344 (0.396)	1.508 (0.440)	50
<i>Friendship Network Centrality</i>						
Network centrality summary index	0.006	-0.671 (0.252)	-0.616 (0.283)	-0.924 (0.342)	-0.873 (0.388)	50
Indegree centrality	0.131	-0.433 (0.256)	-0.444 (0.256)	-0.593 (0.341)	-0.618 (0.334)	46
Outdegree centrality	0.066	-0.529 (0.229)	-0.453 (0.296)	-0.725 (0.343)	-0.631 (0.386)	46
Closeness centrality	0.008	-1.069 (0.411)	-1.020 (0.430)	-1.464 (0.555)	-1.420 (0.585)	46
Eigenvector centrality	0.011	-1.198 (0.504)	-1.111 (0.516)	-1.641 (0.691)	-1.547 (0.712)	46

Heteroscedastic robust standard errors in parentheses.

Table A3. Causal mechanisms: Leader characteristics.

	RI	ITT		IV		<i>n</i>
	<i>p</i> -value	nc	c	nc	c	
<i>Manager Observable Characteristics</i>						
Ability summary index	0.617	-0.094 (0.195)	-0.087 (0.208)	-0.130 (0.261)	-0.123 (0.277)	42
English proficiency (basic)	0.158	0.419 (0.230)	0.400 (0.260)	0.571 (0.321)	0.549 (0.347)	46
English proficiency (advanced)	0.846	-0.090 (0.313)	-0.130 (0.334)	-0.123 (0.416)	-0.178 (0.433)	46
Education attainment	0.395	-0.274 (0.315)	-0.215 (0.315)	-0.378 (0.422)	-0.302 (0.410)	44
Questions solved in raven test	0.648	-0.148 (0.339)	-0.134 (0.341)	-0.206 (0.457)	-0.188 (0.447)	43
Age	0.899	0.467 (3.425)	-0.666 (3.584)	0.635 (4.569)	-0.913 (4.648)	46
Attendance in religious services	0.677	0.093 (0.139)	0.091 (0.150)	0.127 (0.182)	0.125 (0.192)	46
Size of coffee land plot (acres)	0.202	2.850 (2.832)	2.864 (2.593)	3.882 (3.804)	3.930 (3.436)	46
Wealth index (standardized)	0.337	0.559 (0.638)	0.634 (0.645)	0.761 (0.876)	0.870 (0.870)	46
<i>Commitment to Members' Welfare: Behavioral experiments</i>						
Allocation to strangers (DG)	0.149	-0.446 (0.299)	-0.382 (0.296)	-0.607 (0.384)	-0.530 (0.389)	45
Allocation to members (DG)	0.030	0.767 (0.363)	0.921 (0.360)	1.050 (0.494)	1.293 (0.460)	45
Member-stranger allocation (DG)	0.003	1.254 (0.438)	1.360 (0.420)	1.716 (0.569)	1.910 (0.528)	45
Punishment transfer 1 (TPP)	0.232	0.333 (0.187)	0.353 (0.212)	0.457 (0.249)	0.495 (0.283)	45
Punishment transfer 2 (TPP)	0.116	0.448 (0.187)	0.455 (0.204)	0.614 (0.248)	0.640 (0.272)	45
Punishment transfer 3 (TPP)	0.593	0.199 (0.277)	0.183 (0.284)	0.272 (0.374)	0.257 (0.375)	45
Punishment transfer 4 (TPP)	0.462	0.244 (0.319)	0.239 (0.317)	0.335 (0.436)	0.336 (0.420)	45

Heteroscedastic robust standard errors in parentheses.

DG stands for Dictator Game and TPP for Third Party Punishment experiment.

Table A4. Summary statistics.

	Mean	Std. Dev.	Min.	Max.	N
<i>DC Controls</i>					
Original No. of POs	9.780	6.195	4	33	50
Years since APEP began operating	3.660	1.118	2	6	50
Age of DC	2.840	0.997	1	5	50
<i>Other Governance Institutions</i>					
Governance summary index	0	0.451	-0.893	1.401	50
Have a signed constitution	0.800	0.404	0	1	50
Frequency council meetings (on-season)	1.978	1.483	0	5	46
Frequency council meetings (off-season)	1.596	1.393	0	4	47
Council meeting frequency codified in constitution	0.800	0.404	0	1	50
Frequency executives meetings (on-season)	3.043	1.282	1	5	46
Frequency executives meetings (off-season)	2.020	1.317	0	5	50
DC council chairperson term length	2.400	0.915	1	5	45
DC Manager term length	2.275	1.012	0	5	40
Manager attends executive meetings regularly	0.812	0.394	0	1	48
Council representatives can hold other PO positions	0.634	0.488	0	1	41
<i>Leaders' Responsiveness</i>					
Responsiveness summary index	0	0.686	-1.362	1.345	50
Members obtain receipts (M)	1.409	0.379	0.394	1.968	50
Receipts are given to members (R)	0.754	0.229	0.0714	1	50
Members warned: side selling (M)	0.190	0.106	0	0.406	50
Members warned: bad agricultural practices (M)	0.349	0.111	0.111	0.629	50
Manager is very transparent (M)	0.369	0.153	0.114	0.757	50
<i>Members' Cooperation (M)</i>					
Cooperation summary index	0	0.534	-1.033	1.161	50
Members agree to increase commission	0.842	0.113	0.500	1	50
Members paid joining fees	0.538	0.281	0	0.972	50
Members paid annual dues	0.203	0.114	0	0.528	50
Members' contribution: commitment experiment	580.9	329.1	166.7	1477.8	50
Members' share of coffee bulked	0.489	0.186	0.0686	0.879	50
Members planted seedling in past 12 months	0.678	0.150	0.355	0.944	50
Members dries coffee on tarps	0.722	0.235	0.219	1	50
<i>Group Services (E)</i>					
Collective input procurement	0.300	0.463	0	1	50
Loan services to members	0.200	0.404	0	1	50
Pre-season crop purchase	0.184	0.391	0	1	49
Standard coffee prices per kg	1162.5	280.4	450	1693.1	41

(Continued)

Table A4. (Continued)

	Mean	Std. Dev.	Min.	Max.	N
<i>Monitoring</i>					
Monitoring summary index	0	0.677	-1.354	1.754	50
DC council meeting frequency (E)	2.953	2.045	0	11.84	42
Representatives influence over decisions (R)	2.129	0.219	1.550	2.600	50
External auditing (R)	0.452	0.358	0	1.500	50
Internal auditing (R)	0.228	0.110	0.0435	0.517	50
General assembly: reps' attendance (R)	0.787	0.142	0.429	1	50
General assembly: members' attendance (M)	0.566	0.137	0.333	0.889	50
Manager is monitored (R)	0.570	0.194	0.222	0.951	50
Manager is monitored (M)	0.487	0.113	0.222	0.750	50
Manager is accountable (M)	1.160	0.145	0.885	1.448	50
<i>Network Centrality</i>					
Network centrality summary index	0	0.825	-1.902	1.791	50
Manager indegree centrality	14.41	7.265	3	35	46
Manager outdegree centrality	12.30	9.143	0	41	46
Manager closeness centrality	0.654	0.211	0.0323	1	46
Manager eigenvector centrality	0.817	0.189	0.395	1	46
<i>Manager Observable Characteristics</i>					
Manager ability summary index	0	0.626	-1.452	1.269	50
Manager English proficiency (basic)	0.826	0.383	0	1	46
Manager English proficiency (advanced)	0.674	0.474	0	1	46
Manager education attainment	3.159	1.275	1	6	44
Questions solved in raven test	5.674	2.652	1	10	43
Manager's Age	45	10.88	23	76	46
Manager's attendance in religious services	3	0.471	2	4	46
Manager's size of coffee land plot (acres)	3.571	7.340	0.250	50	46
Manager Wealth index	0.761	1.798	-1.279	9.150	46
<i>Behavioral Experiments</i>					
Allocation to strangers (DG)	219.6	132.7	0	500	46
Allocation to members (DG)	446.7	196.1	0	1000	45
Member-stranger allocation (DG)	224.4	202.4	-200	900	45
Punishment transfer 4 (TPP)	0.422	0.499	0	1	45
Punishment transfer 3 (TPP)	0.822	0.387	0	1	45
Punishment transfer 2 (TPP)	0.889	0.318	0	1	45
Punishment transfer 1 (TPP)	0.933	0.252	0	1	45