The Nexus between Gender, Collective Action for Public Goods, and Agriculture

Evidence from Malawi

Nancy McCarthy

Talip Kilic

The World Bank
Development Research Group
Poverty and Inequality Team
March 2014

Abstract

Across the developing world, public goods exert significant impacts on the local rural economy in general and agricultural productivity and welfare outcomes in particular. Economic and social-cultural heterogeneity have, however, long been documented as detrimental to collective capacity to provide public goods. In particular, women are often under-represented in local leadership and decision-making processes, as are young adults and minority ethnic groups. While democratic principles dictate that broad civic engagement by women and other groups could improve the efficiency and effectiveness of local governance and increase public goods provision, the empirical evidence on these hypotheses is scant. This paper develops a theoretical model highlighting the complexity of constructing a “fair” schedule of individual contributions, given heterogeneity in costs and benefits that accrue to people depending, for instance, on their gender, age, ethnicity, and education.

The model demonstrates that representative leadership and broad participation in community organizations can mitigate the negative impacts of heterogeneity on collective capacity to provide public goods. Nationally-representative household survey data from Malawi, combined with geospatial and administrative information, are used to test this hypothesis and estimate the relationship between collective capacity for public goods provision and community median estimates of maize yields and household consumption expenditures per capita. The analysis shows that similarities between the leadership and the general population, in terms of gender and age, and active participation by women and young adults in community groups alleviate the negative effects of heterogeneity and increase collective capacity, which in turn improves agricultural productivity and welfare.

This paper is a product of the Poverty and Inequality Team, Development Research Group. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at <http://econ.worldbank.org>. The authors may be contacted at tkilic@worldbank.org.

The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

The Nexus between Gender, Collective Action for Public Goods, and Agriculture: Evidence from Malawi¹

Nancy McCarthy[†] and Talip Kilic[‡]

JEL Codes: C26, D71, H41, O13, Q12.

Keywords: Public Goods, Collective Action, Gender, Agriculture, Sub-Saharan Africa, Malawi.

Sector Board: Agriculture and Rural Development (ARD)

¹[†] Senior Researcher, LEAD Analytics, Inc., Washington, DC (nmccarthy@leadanalyticsinc.com); [‡] Research Economist, Development Research Group, The World Bank, Washington, DC (tkilic@worldbank.org). The authors thank Laura Dick for excellent research assistance, and Markus Goldstein, Calogero Carletto, Paul Winters, Alan de Brauw, and participants of the World Bank-IFAD Gender and Agriculture in Sub-Saharan Africa workshop (September 26-27, 2013; Rome, Italy) for their comments. While the findings, interpretations, and conclusions expressed in this work fully reflect the views of LEAD Analytics, they do not necessarily reflect the views of the World Bank, its Board of Executive Directors, or the governments they represent.

1. Introduction

In the absence of a formal government, public goods and services in many rural areas of the developing world are often provided collectively, reliant on voluntary participation and contribution by community members. These public goods and services can have substantial impacts on agricultural production and productivity, with concomitant impacts on household welfare and livelihood strategies. However, impacts are likely to differ across different community members, by gender, age, ethnicity, education and wealth. For instance, wealthy, older male-headed households, who are more active in cash crop production or livestock rearing, may gain relatively more from improved transportation infrastructure or a community livestock crush than poorer female-headed households or households headed by young adults and/or marginalized ethnicities. Even women and men within the same household may receive different benefits from the same public good. Men and women, as well as old and young, may also face different opportunity costs in contributing to the public good. Such economic and socio-cultural differences have long been posited as key factors that limit local capacity to provide public goods. Economic heterogeneity leads to divergent costs and benefits across members, and complicates the negotiation of agreements underlying communal contributions. Heterogeneity in socio-cultural norms and differing degrees of trust across different demographic and ethnic strata can further increase costs associated with negotiating, monitoring and enforcing agreements.

There are a number of ways to mitigate the costs imposed by heterogeneity on collective provision of public goods. More representative local leadership and increased participation in local organizations by women and other diverse community members can give “voice” to divergent interests and allow for more transparent and inclusive negotiations. To the extent that such negotiations result in agreements that are seen as fair in the sense that they accommodate different costs and benefits among community members, these agreements should subsequently be easier to monitor and enforce. This paper focuses on the extent to which communities, through broad civic engagement in local organizations and more representative leadership, can ameliorate the negative impacts of economic and socio-cultural heterogeneity on collective capacity to provide public goods and improve welfare in rural Malawi. While the emphasis is on women’s representation in leadership and active participation in local organizations, to properly analyze the impacts of heterogeneity, we also include other dimensions of diversity including age, education, ethnicity and wealth.

This research sits at the intersection of two strands of literature. The first encompasses game-theoretic models of local public goods provision that explicitly account for the impacts of heterogeneity, and the second is a multi-disciplinary body of research that focuses on participatory development. The first starts from the “problem” of public good provision in the face of heterogeneity, and arrives at the conclusion that in most cases, heterogeneity has negative impacts on the ability to provide public goods through collective action. Although the empirical analyses that test this hypothesis find wide support, this literature rarely focuses on the

mechanisms that might be used to ameliorate negative impacts of heterogeneity. The second strand of literature, on the other hand, starts from the premise that greater civic engagement facilitates public good provision, in part because such engagement alleviates negative effects of heterogeneity. Within this strand, empirical studies document factors that facilitate or inhibit collective action outcomes in community-driven development and social fund projects in developing countries, with some papers specifically focusing on the role of women's leadership representation and voice. A more limited set of papers attempts to identify the effect on welfare of local public good provision through collective action.

This paper ties together both strands of literature by first developing a stylized theoretical model of "community welfare optimization", and then by testing hypotheses stemming from the model empirically. We build on the insights from the game theoretic literature and model a community seeking to maximize social welfare but subject to transactions costs of acting collectively. In particular, we posit that negotiating, monitoring and enforcement costs are a function of the schedule of individual members' agreed contributions. Schedules that fully reflect underlying heterogeneity are seen as "fair", and lead to lower monitoring and enforcement costs; however, negotiation costs are higher than agreeing on more simple schedules, such as equal contributions by all. While highly stylized, maximizing over community welfare as a function of transactions costs enables us to explicitly highlight informational requirements necessary to construct such a fair schedule in the face of heterogeneity, which is often obscured in extant game-theoretic models.

We empirically test the theoretical model by utilizing the unique data solicited through the Household and Community Questionnaires of the nationally-representative Malawi Third Integrated Household Survey (IHS3) 2010/11 and linking the geo-referenced HIS3 household and community data to geospatial and administrative data sources. Subsequently, we (i) formulate a composite index, based on principal components analysis, of community collective capacity for public good provision, (ii) regress the index as a function of a rich set of attributes that corresponds to the theoretical model, and (iii) examine the effects of collective capacity for public good provision on community median estimates of maize yield and household real annual consumption expenditures per capita in a Two-Stage Least Squares Framework that takes into account the non-random nature of the explanatory variable of interest. Our data-intensive exercise is conducted at the community level, limited to the rural enumeration areas visited as part of the IHS3. By relying on the IHS3, and bringing in data from other geospatial and administrative sources, the paper contributes to the otherwise limited empirical analyses based on nationally representative data with large sample sizes. To the best of our knowledge, no large-sample analyses have been able to test whether community leadership that reflects underlying community heterogeneity improves public good provision, although there is some case study evidence to support that hypothesis.

The choice to focus on Malawi is not only driven by the availability of rich data but also by the unique process of decentralization that the country has been undergoing since the adoption of the Local Government Act and the Decentralization Policy in 1998. The legal framework outlines the structure, roles and responsibilities of District Councils (DC), and allows the DCs to facilitate the establishment of sub-district structures, such as Village Development Committees (Chinsinga, 2008).² Guidelines for the establishment of sub-district structures state that the VDC should be a representative body with at least four women representatives, and that the Group Village Head should supervise, and not chair, the VDC (Chiweza, 2010). Anecdotal evidence, however, suggests that there are spatial differences in the representation of women in the VDC and that Village Heads often chair the VDCs. Furthermore, the roles, responsibilities and interactions of various sub-district entities remain unclear to most citizens (Chiweza 2010).

As part of the decentralization process, it was envisaged that communities would draw up development project proposals and seek funding for those projects. Currently, many communities do seek funding from external sources, including those established through the decentralization framework, e.g. Local Development Funds, but also from other sources such as the Constituency Development Funds under the direct control of Members of Parliament, and non-governmental organizations (NGOs). Nonetheless, communities also still rely on member-provided labor, cash and in-kind materials to provide public goods as well, so that the system is a hybrid between more formally-sourced, externally financed public goods provision and locally-based provision financed by voluntary contributions.

Four key findings emerge from the empirical analysis. First, similarity between the local leadership and the general population in terms of gender and age, and shares of women and individuals under the age of 30 in the total membership across non-directly productive local organizations are positively associated with collective action capacity for public good provision. Second, community heterogeneity in terms of ethnicity, education and wealth, exert statistically insignificant effects on collective capacity; a result that is consistent with the hypothesis that more representative leadership and broad civic engagement in local organizations can be successful in alleviating the negative effects of heterogeneity. Third, collective capacity is positively correlated with favorable rainfall and agro-ecological zone placement. This result is in line with the hypothesis that more promising average agro-climatic conditions increase returns to local public goods and thus favor greater collective action capacity. Greater climate variability also leads to greater collective capacity, consistent with limited individual/household-based strategies to limit exposure to climate risks. Fourth, aligned with the assertion that public goods provided through voluntary participation and contribution by community members can have substantial impacts on productivity and welfare outcomes, the results suggest that collective

² The decentralization process has been stalled in terms of elections for district councilors; in particular, a new round of elections to be held in 2005 has yet to occur (Chinsinga, 2008; Chiweza, 2010). This vacuum has been filled by parallel structures in some instances, and overall, the exact role of sub-district structures remains confusing for citizens (Chiweza, 2010).

action capacity for public good provision exerts positive effects on community median estimates of maize yields and household real annual consumption expenditures per capita.

The remainder of the paper is organized as follows. Section 2 provides a succinct literature review, followed by the presentation of a theoretical model of public good provision, leadership composition and voice. Section 4 describes the data and the estimation strategy. Section 5 discusses the core results and the findings from the robustness analyses, and Section 6 concludes.

2. Literature Review

There is a wide body of theoretical and empirical literature examining the factors that aide or hinder collective action in the provision of public goods and/or the management of communal resources (cf. Olson, 1965; Ostrom, 1990; Bromley, 1992; Sandler 1992; Baland and Platteau, 1996). Generally, groups of individuals wishing to benefit from local public goods create an institutional framework to internalize externalities; externalities that otherwise lead to under-provision of public goods when individuals only optimize over their own private costs and benefits. Both the theoretical and empirical literature identify a number of factors important for determining the capacity of groups to collectively provide public goods. The most widely studied factor emerges as “heterogeneity”. However, community members can diverge along many dimensions, including age, gender, ethnicity, education, wealth, and income. Differences across these dimensions can have different impacts on the ability to provide public goods. The literature broadly groups these dimensions into two; economic heterogeneity and socio-cultural heterogeneity.

Economic heterogeneity – often measured in terms of different wealth levels – is often posited to reduce the capacity to collectively provide public goods because it requires reconciling divergent benefits and costs associated with providing such goods (Ostrom, 1990; Baland and Platteau, 1997). Dayton-Johnson (2000) develops a non-cooperative game theoretic model that captures disincentives to contribute to the public good when contributions are the same for all members regardless of heterogeneity in benefits. Due to lower transactions costs of implementing a flat contribution versus proportional contribution schedule, the flat contribution schedule may still be chosen. Others have argued that economic heterogeneity favors collective action; for instance one can construct cases where the wealthy find it in their best interest to provide certain public goods whether or not others contribute (Olson, 1965; Sandler, 1992; Bardhan et al., 2007).³ In the latter case, private returns to the wealthiest are sufficient to induce provision, even if others can benefit freely.

³ Whether or not there are farmers’ wealthy enough to finance the public good is not necessarily captured by measures of dispersion around the mean, rather it has to do with the tail of the distribution. Nonetheless, the two are generally correlated, and very often conflated in the literature.

Socio-cultural heterogeneity may lead to divergent incentives to provide specific public goods, and may make negotiating and enforcing an agreement more difficult to the extent that trust amongst groups is lower than within groups, and also given different underlying “norms” (Alesina and La Ferrara, 2000; Ostrom and Walker, 2003; Ruttan, 2008). Since such groups often rely on “social suasion” to make and enforce agreements, socio-cultural diversity can significantly weaken the impact of social sanctions (Miguel and Gugerty, 2005). Socio-cultural heterogeneity, on the other hand, has also been claimed to mitigate “institutional inertia” associated with more isolated, homogeneous communities (Begossi, 1998), and to increase the capacity to provide public goods by bringing to the table different groups with distinct skills and experiences (Alesina and La Ferrara, 2005 and references cited therein).

Empirical evidence on the impact of economic heterogeneity is mixed, but largely favors the hypothesis that it decreases the capacity to undertake collective action and thus provide public goods (Johnson and Libecap, 1982; Varughese and Ostrom, 2001, Bardhan, 2000; Gebremedhin et al., 2004; McCarthy, 2004; Andersson and Agrawal, 2009). Studies that find positive impacts tend to find these under particular circumstances of economies of scale with relatively high returns to the wealthiest, as discussed above (Ruttan, 2008; Naidu, 2009). Results from empirical studies that look at the impact of socio-cultural heterogeneity vary as well, but generally support the hypothesis that heterogeneity reduces collective action capacity to provide public goods. In India, Baland et al. (2010) and Bardhan and Dayton-Johnson (2007) find evidence that heterogeneity in caste hinders collective action to manage firewood extraction and irrigation, respectively. La Ferrara (2002) finds evidence that ethnic fragmentation reduces the effectiveness of production cooperatives, and shows that more heterogeneous group choose a simple payment scheme versus a schedule of task-specific payments. The author argues that heterogeneity makes the transparency and negotiation costs of the simple payment scheme relatively more attractive, even if otherwise less efficient than the per task schedule.

Heterogeneity, however, can be managed. Collier (2001) argues that democracies are better able to manage ethnic diversity to the extent that different ethnicities feel they have at least some representation. A number of researchers have empirically looked at the impact of gender and other demographic characteristics of both participants and leaders on mitigating the negative impacts of heterogeneity (Agarwal, 2009; Pandolfelli et al., 2008; Godquin and Quisumbing, 2008). For instance, in forestry cases in India and Nepal, Agarwal (2009) finds that having a high proportion of women on the executive committee of the forest management group is correlated with a greater improvement in forest condition; this effect is even greater if these women are older. She suggests, following Westermann et al. (2005), that women are better at protection, compliance, and at fostering cooperation among other women, and indeed finds that compliance rates are higher among women when women are on the executive committee, perhaps because they feel they have some ownership over the process and rules.

Barham and Chitemi (2009), in examining collective action among farmer smallholder groups in Tanzania find that female-dominated groups (defined both in terms of leadership and membership) were in fact disadvantaged in regards to marketing their produce, as women faced a time disadvantage in seeking new markets as well as reduced access to non-local socio-political networks. Pandolfelli et al. (2008) explain this and related findings as related to women's reduced likelihood to be tied into traditional information networks. However, in many cases women themselves rely on information provided by collective action institutions, so women's participation is necessary to voice demand for relevant information. The authors stress the different complementary roles played by men and women in collective action: for example, men specialize in physical guarding, while women apply their efforts to social pressures. This implies that collective action groups with different gender compositions behave differently: women-only groups are strong on participatory practices, men-only on setting rules, but that it is only mixed groups that truly prompt community-wide actions, increasing the effectiveness of collective action (c.f. Were et al 2008).

A number of other variables are associated with the capacity to engage in collective action. Outside of heterogeneity, group size is the most well-studied. Early studies focused on increasing transaction costs as group size increases, implying reduced collective action (Olson, 1965; Ostrom, 1990). On the other hand, for public goods with high fixed costs (e.g. irrigation infrastructure), small groups may simply be unable to afford provision (de Janvry et al, 1998). Additionally, if the public good exhibits increasing returns to scale over some range, larger groups will be more able to take advantage of these returns (Karaivanov, 2009). Where community members have greater options to substitute for collective action by private initiative, the ability to act collectively is often reduced (Ostrom, 1990). For instance, where members have greater options for working outside of the community, collective maintenance of irrigation infrastructure, was found to be lower (Bardhan, 1993; Dayton-Johnson and Bardhan, 2002). Supra-local institutions may also either substitute or compliment local collective action, e.g. public works and social safety net programs.

Finally, there is a separate but related literature on decentralization, devolution and participatory development, which we touch on cursorily.⁴ Mansuri and Rao (2012) note that dissatisfaction with the highly centralized development projects of the 1970s and 1980s led to a shift towards “local participatory development”. Involvement of local communities in a wide range of development activities (including local public goods provision) was posited to result in better designed and more effectively implemented projects, as well as broad-based, inclusive economic development (Sen, 2009). The empirical evidence suggests that the success of local participatory development projects hinge critically on whether the project actually “induces” sufficient local

⁴ Mansuri and Rao (2012) and Dreze and Sen (2002) provide detailed reviews of the history, theoretical underpinnings, and empirical analyses of participatory development projects.

participation. The case study literature highlights two important factors in the ability to induce local participatory development, education levels of both community members and leaders, and the degree of isolation of the community (Mansuri and Rao, 2012). Dreze and Sen (2002, chapter 10) also note education levels as being important in explaining differences in progress towards more inclusive participatory development between Indian states. However, there is a dearth of information from large-sample empirical analysis, particularly evidence of the link between broad-based participation and representative leadership on collective action and public goods provision, and subsequent impacts on poverty and other indicators of household well-being.

3. Theoretical Model of Public Goods, Leadership Composition, and Participation

As with private goods, the optimal amount of public goods to produce will be a function of marginal benefits and marginal costs. However, to reach the social optimum there will be costs associated with internalizing externalities, e.g. costs of negotiating, monitoring and enforcing agreements. In reaching an agreement, the group needs to choose how contributions will be allocated; for simplicity, below we compare two allocation mechanisms: 1) equal for all members, and 2) a schedule that reflects the net marginal benefits accruing to each group member. Equal contributions have the great advantage of being simple and transparent. Negotiating a schedule of contributions can be costly and might still leave some feeling the schedule is unfair given asymmetric information regarding individual-specific costs and benefits. To be able to handle heterogeneity, it is hypothesized that a diverse set of community leaders for instance in terms of gender, age, education, ethnicity and wealth, will be better able to facilitate negotiations over a schedule of “fair” contributions. When diverse members also actively participate in community life, their concerns will be voiced more strongly, so that leadership can adequately take these concerns into account when deciding on what public goods to provide and how to allocate different contributions, given the distribution of subsequent benefits.

To highlight the informational requirements needed to develop a schedule of contributions that allocates contribution costs in proportion to benefits, we calculate the each individual’s contribution that results from the social optimizer’s problem, restricting attention to a pure private good as follows:

$$\max_{X^{PG}} SW(X^{PG}) = \sum_i (V_i(X^{PG}) - c_x X^{PG}) \quad [1]$$

Where SW is the social welfare, X^{PG} is the sum of all members contributions, x_i^{PG} ; $V_i(X^{PG})$ is the value to the i^{th} member from total contributions, and c_i^x is the unit cost borne by the i^{th} member. To simplify matters, we initially assume that c_x is constant across all members. The resulting

first-order condition for the social optimizer is as shown, where we separate the marginal value to the i^{th} member out to facilitate comparison with the non-cooperative outcome:

$$\frac{\partial \text{SW}(X^{PG})}{\partial X^{PG}} = \frac{\partial V_i(X^{PG})}{\partial X^{PG}} - c_x + \frac{\partial \sum_{j \neq i} V_j(X^{PG})}{\partial X^{PG}} \quad [2]$$

The individual's first order conditions would be the same as the first two terms in equation [2], but would not include the additional value to other members. Thus, the first-order condition for the individual would be lower than for the social optimizer, and under-provision would result. We can still see that even for the individual, the total amount provided will depend on marginal costs and private marginal benefits. Even under non-cooperation, the amount provided will be greater the higher are (private) benefits relative to costs.

As is well known, the solution to the social optimizers problem yields an optimal total amount of the pure public good, X^{PG*} , but does not reveal how the burden is to be shared amongst members. If members are homogeneous – they receive the same marginal benefit and face the same marginal costs – then it is natural to think about splitting total provision by the total number of members. An even split means that all members would still have the same net absolute gains when both their value function and marginal costs are the same:

$$V_i(X^{PG}) - c_{xi} x_i^{PG} = V_j(X^{PG}) - c_{xj} x_j^{PG}, \forall i, j.$$

With heterogeneity, to determine each individual's equitable contribution, we specify the net value equation for the representative “other” community member using the mean value for all non- i^{th} individuals as:

$$\bar{V}_{j \neq i}(X^{PG}) - \bar{c}_x \bar{x}_{j \neq i}^{PG}$$

Noting that

$$\bar{x}_{j \neq i}^{PG} = \frac{(X^{PG} - x_i^{PG})}{(n-1)},$$

we calculate an equitable contribution value for the i^{th} member by setting

$$V_i(X^{PG}) - c_{xi} x_i^{PG} = \bar{V}_{j \neq i}(X^{PG}) - \bar{c}_{xj \neq i} \frac{(X^{PG} - x_i^{PG})}{(n-1)}.$$

Re-arranging the terms gives us the equitable x_i^{PG} as a function of the difference in marginal benefits minus costs, deflated by cost terms:

$$x_i^{PG} = \frac{V_i(X^{PG}) - \bar{V}_{j \neq i}(X^{PG})}{\left(c_{xi} + \frac{\bar{c}_{xj \neq i}}{(n-1)}\right)} + \frac{\bar{c}_{xj \neq i} X_j^{PG}}{(n-1) \left(c_{xi} + \frac{\bar{c}_{xj \neq i}}{(n-1)}\right)} \quad [3]$$

When $V_i(X^{PG}) = \bar{V}_{j \neq i}(X^{PG})$ and $c_{xi} = \bar{c}_{i \neq j}$ for all i , then $x_i^{PG} = \frac{X^{PG}}{n}$, an equal split. x_i^{PG} increases as the difference in the value of the public good vis-à-vis other members increases, and to the extent that the i^{th} member has both an absolute and relative cost advantage. Note that x_i^{PG} can be negative if the value difference is sufficiently negative. Though a very simple representation, the model highlights how complicated it can become to precisely determine a schedule of equitable contributions to the public good; leaders constructing such a schedule will have to know each members' value and cost functions in order to make adjustments from the equal split rule that are seen as fair by all community members.

The social welfare maximization problem assumes zero costs of collective action. Given the costs of collective action discussed above, we can write the community-members' optimization problem as follows:

$$\max_{X^{PG}} CW(X^{PG}) = V_i(X^{PG}) - c_{X^{PG}} \left\{ 1 + (TC(H); Z^L, Z^{Comm}) \right\} X^{PG} \quad [4]$$

where CW is community welfare, $c_{X^{PG}}$ are marginal costs of collective action, and TC are transactions costs specific to negotiating, monitoring and enforcing agreements. When transactions costs are zero, the community welfare optimization problem reduces to the social optimizer's problem. H is heterogeneity across members in terms of different public goods' valuation and costs of contributing to the public good. Finally, Z^L is a vector of characteristics of community leadership, and Z^{Comm} is a vector of other community characteristics affecting the cost of providing the public good. In what follows, we consider all factors that lower marginal costs of collective action (TC, Z^L, Z^{Comm}) as capturing "cooperative capacity." Maximization provides an optimal total amount of the good to provide, X^{PG} , given the outcome of negotiations over allocation of costs and benefits, and given subsequent monitoring and enforcement costs.

Specifically, we posit that

$$\frac{\partial c_{X^{PG}}}{\partial TC} \frac{\partial TC}{\partial H} > 0;$$

so that increased heterogeneity leads to greater marginal costs associated with negotiating, monitoring and enforcing agreements. All else equal, the total amount provided declines as heterogeneity increases. We consider heterogeneity in ethnicity, education and landholdings (as a proxy for wealth) as part of the empirical test of the model. Since communities are quite similar in terms of gender and age distributions, it is not possible to control for the impact of gender and age heterogeneity directly. In addition, as noted in the literature review, certain measures of diversity may in fact decrease costs of collective action. For instance, ethnic diversity may provide a greater range of experiences on which to base collective action. Our null hypothesis is that heterogeneity nonetheless increases costs of collective action, with the alternative being that some forms of heterogeneity may decrease such costs.

A community can mitigate the negative impacts of heterogeneity by accommodating different interests. We hypothesize that when leader characteristics are similar to the general population, negotiating costs in particular will be lower, as leaders will be better able to take into account underlying heterogeneity within the community, assessing the distribution of costs and benefits that will accrue across different members in the community. For instance, where leadership is roughly balanced between men and women, mirroring the underlying gender composition of the community population, we expect that negotiating, monitoring and enforcement agreements will be lower, as different gender-based interests will be more likely to be accommodated. Similarly, we would expect that having leaders whose age reflects the distribution of adult ages in the population will be more effective at accommodating age-based difference in interests. As part of our empirical application, we also provide measures of similarity between leaders and general population in terms of ethnicity and education, again with the hypothesis that the costs of reaching and enforcing an agreement will decrease with increasing similarity.

While some of the literature does document the benefits to having diverse community leaders, to the best of our knowledge, this is the first empirical application that examines the role of similarity between the leadership and the general population in facilitating collective action outcomes by using similarity indices whose construction is detailed in the subsequent section. Another mechanism by which heterogeneity of interests can be more easily accommodated is active participation of members with different interests in community life. Participating in different local organizations is one mechanism by which diverse groups can express their “voice.” In this light, we claim that broad civic engagement across gender and age categories in local organizations reduces costs associated with accommodating heterogeneous interests, hence costs of undertaking collective action.

A number of other leadership and community level characteristics are also posited to affect collective action. In terms of leadership, absolute levels of education and age are hypothesized to have a separate impact on costs of collective action. Greater education levels should be associated with greater managerial and problem-solving skills. Older members often have greater

network connections and a wider range of experiences within and outside the community that can expand capacity to promote collective action. Following the literature review, the size of the community can too affect collective action. While very small communities may face high per-unit costs, larger communities face more difficulties in internalizing externalities, even where community members are homogeneous.⁵

In addition to transactions costs of collective action, there are factors that determine the benefits and costs to undertake collective action. Though not explicit in the theoretical model, people can substitute away from local public goods if collective provision is too costly (McCarthy and Essam, 2009; Alesina and La Ferrara, 2005). This makes it difficult to sign many factors that are likely to affect the benefits and costs of public goods as well as potential substitutes. “Exit options”, often proxied by temporary or permanent out-migration have been argued to reduce cooperative capacity by increasing opportunity costs of allocating resources to public goods provision (Dayton-Johnson and Bardhan, 2002). On the other hand, earnings from migration could be re-invested on-farm, increasing public goods that have positive spillover benefits on yields and/or marketing (planting windbreaks, investing in water and soil conservation management structures, maintaining bridges and roads). While more favorable agro-ecological and climate characteristics should increase returns to public goods provision, the effect of greater production risks is not clear and depends on whether public goods are a cost-effective way to manage greater risks vis-à-vis purely private management strategies (McCarthy, 2004). Access to infrastructure and services not provided by local collective action, such as credit and agricultural extension, can also increase or decrease collective capacity, depending on whether such infrastructure and services act as substitutes or compliments to the provision of local public goods.

To summarize, the model suggest six categories of variables that could affect collective action capacity: 1) underlying community heterogeneity, 2) representation and voice, which mitigate negative impacts of heterogeneity, 3) leadership capacity, 4) community characteristics affecting transactions cost of collective action, 5) agro-ecological and climatic conditions, and 6) infrastructure and services, not provided by collective action. In what follows, we provide empirical tests of the hypotheses stemming from the theoretical model. Our primary goal is to estimate a multivariate model of community collective capacity to provide local public goods as a function of variables that are directly mapped to the factors underlying collective capacity identified above. The secondary objective of our analysis is to look at the effect of community collective capacity to provide local public goods on proxies of community agricultural productivity (median maize yield) and welfare (median household real annual consumption expenditures per capita). Given the concerns of unobserved heterogeneity that may jointly determine collective action and productivity/welfare outcomes at the community level, we

⁵ In a standard public goods provision model, the extent of under-provision increases as membership increases, even when members are homogeneous (Sandler, 1992).

attempt to fulfill our secondary objective within an instrumental variable regression framework. The subsequent section describes our data, provides definitions for indices of community collective capacity and measures of diversity and similarity. This section concludes with descriptive statistics, and details the estimation strategy.

4. Empirical Approach

4.1. Data

Our analysis is data-intensive, capitalizes on the multi-topic nature of the Third Integrated Household Survey (IHS3) 2010/11 survey instruments and takes advantage of administrative and geospatial data that could be linked with the communities of interest. The IHS3 was implemented from March 2010 to March 2011 by the Malawi National Statistical Office, with support from the World Bank Living Standards Measurement Study - Integrated Surveys on Agriculture (LSMS-ISA) project.⁶ The IHS3 enumeration areas (EAs) and households were selected as part of a two-stage cluster sampling design with the EAs, in each survey stratum, selected in the first stage with probability proportional to the household count from the 2008 Population and Housing Census. Households were then selected at random in the second-stage in each sampled EA following a full household listing. The 31 IHS3 strata correspond to all 27 rural districts, with the exception of Likoma Island, and 4 urban cities, namely Lilongwe, Blantyre, Zomba, and Mzuzu. The total sample is 12,271 households in 768 EAs (approximately 16 households per EA) that is representative at the national-, urban/rural-, regional, and district levels. For the purposes of our analysis we focus on the rural sample, consisting of 628 EAs.

The IHS3 instruments included Household, Agriculture, Fishery, and Community Questionnaires. All sample households were geo-referenced and administered the multi-topic Household Questionnaire that collected individual-disaggregated information on demographics, education, health, wage employment, nonfarm enterprises, anthropometrics, and control of income from non-farm income sources, as well as data on housing, food consumption, food and non-food expenditures, food security, and durable and agricultural asset ownership, among other topics. The data allow for the computation of a comprehensive household consumption aggregate, which in turn enables us to calculate community level median estimate of household real annual consumption expenditures per capita as one of our measures of welfare.⁷

⁶ The authors include the point person for the World Bank technical assistance towards the design and implementation of the IHS3 under the LSMS-ISA initiative, which is a household survey program established by a grant from the Bill and Melinda Gates Foundation to provide financial and technical support to governments in sub-Saharan Africa in the design and implementation of nationally-representative multi-topic panel household surveys with a strong focus on agriculture (www.worldbank.org/lms-isa). The IHS3 data and documentation are publicly available through the LSMS website (www.worldbank.org/lms).

⁷ The analysis uses the official consumption aggregate that was jointly calculated by the NSO and the World Bank, as documented in Chapter 13 and Appendix B of the NSO IHS3 household socio-economic characteristics report.

The sample households that were involved in agricultural activities (through ownership and/or cultivation of land, and/or ownership of livestock) were administered the Agriculture Questionnaire. Handheld global positioning system (GPS)-based locations and land areas of the plots were recorded. The Agriculture Questionnaire also solicited information on physical characteristics, labor and non-labor input use, and crop cultivation and production at the plot level, separately for the reference rainy and dry seasons.⁸ Depending on the timing of the interview, the reference rainy season could have been 2008/09 or 2009/10, while the reference dry season could have been 2009 or 2010. We rely on the observed plot level distribution of rainy season maize yield (i.e. kilogram-equivalent production per hectare), and compute the median maize yield in each community as the second outcome variable.⁹

Pertinent to our research, the Community Questionnaire was administered to a focus group concerning a village, a group of villages or an urban ward existing in each IHS3 EA. In rural Malawi, a typical EA corresponds roughly to 2 to 3 villages, approximately 250 households. During the IHS3 field work, the boundaries of each EA were clearly established through the use of maps that were produced by the NSO Cartography Department. The leader of the mobile survey team that was assigned to a given EA was responsible for administering the Community Questionnaire. The team leader was instructed to form a focus group that was composed of 5 to 15 long-term knowledgeable residents of the community and that was diverse in terms of sex, age, religion and ethnicity. The focus group members typically included the village chief(s) and the advisors to the village chief(s), a subset of members of the village development committee (VDC) or the area development committee (ADC), the local school headmaster and/or teacher, the health worker(s), the agricultural extension officer, the leaders of religious and political entities, the local merchants, the leaders and members of community-based organizations/committees, and members of community policing.

Table 1 presents descriptive statistics on the composition of the focus groups as captured in the data. The size of the focus group ranges from four to 12 with a mean of 8.5. At the community level, 91 percent of the community focus groups had at least one of village chief, group village

The report could be accessed publically from http://siteresources.worldbank.org/INTLSMS/Resources/3358986-1233781970982/5800988-1271185595871/IHS3_Report.pdf.

⁸ A plot was defined as a continuous piece of land on which a unique crop or a mixture of crops is grown, under a uniform, consistent crop management system, not split by a path of more than one meter in width. Plot boundaries were defined in accordance with the crops grown and the operator. The IHS3 identified 18,917 plots that were reported to have been owned and/or cultivated during the reference rainy season.

⁹ The standard approach to collecting production data at the plot-crop level is to allow for non-standard production measurement units to be specified. To obtain the kilogram-equivalent production values, the authors used a conversion factor database that was constructed following the IHS3 fieldwork through a market survey that visited 22 markets throughout the country, chosen on the basis of spatial crop marketing patterns. This database provides kilogram-equivalent weight estimates for non-standard measurement unit-crop combinations observed in the data. For maize specifically, the Agriculture Questionnaire solicited some information on the state of the crop, specifically whether reported production was shelled or unshelled. The conversion factor database also provides a shelled (i.e. grain) equivalent conversion for unshelled maize production so that all production could be expressed in shelled-equivalent terms.

chief, or traditional authority as a respondent. 98 percent had either one of the aforementioned leaders, a village chief counselor, or ADC/VDC member as a respondent. The leaders/members of religious entities, the leaders/members of community-based organizations/committees, and the business women/men made up 10.7, 9.9 and 8.3 percent of the focus group members, respectively. The roles of the rest of the 22 percent of all respondents were local headmaster/teacher (6.8 percent), health worker (3.5 percent), police (1.9 percent), agricultural extension staff (1.3 percent), political party leader/member (0.5 percent), and other (8.0 percent). 28 percent of the focus group respondents were women, with an average age of 45 years. On average, the respondents reported to have spent 79 percent of their lifetime in the community and 44 percent of them had at least a primary school diploma. The ethnic distribution of the respondents is in line with the dynamics of the rural Malawian general population. Given the findings reported in Table 1, we assume the focus group to represent leadership within the community.

The last set of variables informing our analyses originates from geospatial and administrative data that are mapped to the 628 rural EAs of interest. The geospatial data originate from a public data set that is available alongside the IHS3 raw data and that is at the household level, obtained by linking GPS-based household locations with public global geospatial data sets of varying degrees of resolution.¹⁰ Given the resolution of the available global geospatial data, the geospatial variables that are derived at the household level exhibit limited variation within the EAs, if any. Nonetheless, community level geospatial variables are constructed by taking the median values of their household level counterparts in each community. Finally, four administrative variables are featured in the analysis, three of which are defined at the district level; these include the 2007/08 agricultural season per household measure of metric tons of subsidized fertilizer delivered to the district, the 2008/09 agricultural season per household measure of the total wages paid as part of the Malawi Social Action Fund (MASAF) that serves as a social safety net and a public works program, the 2008/09 agricultural season number of agricultural development projects operating in the district.¹¹ The fourth administrative variable has sub-district specificity and is defined at the extension planning area (EPA) level. This variable captures the number of farmer organizations operating in the EPA.¹² Table A1 in the Appendix maps all variables featured in the analysis to their sources.

¹⁰ The list of derived geospatial variables that accompany the IHS3 unit record data could be obtained here: <http://siteresources.worldbank.org/INTLSMS/Resources/3358986-1233781970982/5800988-1271185595871/Geovariables.Description.pdf>.

¹¹ This administrative data was compiled and shared with us by the EPIC project, led by the Agriculture and Development Economics Division (ESA) of the United Nations Food and Agriculture Organization. The data was collected from the Malawi Ministry of Agriculture and Food Security. The normalization of the administrative data by the number of households is based on the district-level estimates of household population that is obtained by aggregating the IHS3 household sampling weights at the district-level.

¹² This data was provided by the Farmers Union of Malawi. An EPA is the basic operational and administrative unit for the extension system.

4.2. Collective Action Capacity Index

To apply the model empirically, we first develop an index of collective capacity that captures the ability to provide public goods. Since our ultimate goal is to determine whether collective capacity to provide local public goods has positive impacts on agricultural productivity and welfare, we create an index based on a range of activities tied to public good provision within the community instead of focusing on any one public good in particular (McCarthy et al., 2004; Narayan and Pritchett, 1999). In large part, this is due to the fact that we have information on 17 types of public goods, all of which may affect communal agricultural performance and welfare directly or indirectly, and it is not feasible to take into account the endogeneity associated with 17 different dimensions.

The collective capacity index is based on the raw data collected through the IHS3 Community Questionnaire Module CH: Community Needs, Actions and Achievements. In Module CH, the focus group was probed regarding the community's need for 17 distinct public goods during the past five years, and was given an option to specify any other public good that was not captured by the list and that the community may have needed. The list of public goods were calibrated through piloting ahead of the actual field work implementation, and included (i) road: construction, (ii) road: maintenance, (iii) bridge: construction, (iv) bridge: maintenance, (v) primary school: construction, (vi) primary school: maintenance/ improvement, (vii) secondary school: construction, (viii) secondary school: maintenance/ improvement, (ix) health center/ clinic/ dispensary: construction, (x) health center/ clinic/ dispensary: maintenance/ improvement, (xi) piped water/ boreholes/ wells: construction, (xii) piped water/ boreholes/ wells: maintenance/ improvement, (xiii) orphanage: construction/ maintenance/ improvement, (xiv) maize mill: construction/ maintenance/ improvement, (xv) public transportation: initiation/ improvement, (xvi) agricultural/ fishery/ livestock extension services, and (xvii) law enforcement: initiation/ improvement. We were able to classify into one of the 17 categories all other public goods that were additionally specified by the focus group.

For each public good, the module included a series of questions to capture the specific actions that the community may have taken to address the need. The actions that Module CH inquired about included (1) mobilizing manpower among community members, (2) mobilizing financial resources among community members, and (3) submitting a proposal to be funded through various external mechanisms, namely MASAF/Local Development Fund, Traditional Authority, District Assembly/Commissioner, Member of Parliament, and non-governmental/religious/international organization. The focus group was also given an option to specify any other action that the community may have taken. Following these questions, the focus group was asked whether a particular public good need was addressed/being addressed, and *conditional* on a positive response, provided information on whether the community

members contributed financial resources (money and in-kind support) and/or supplied manual labor of any kind during the period that the need was met.

The first step in the construction of the collective capacity index was to classify the 17 public goods into five broad *domains* namely (1) infrastructure (i through iv, and xv), (2) education (v through viii, and xiii), (3) health (ix through xii), (4) agriculture (xiv and xvi), and (5) law (xvii). The second step was to classify all types of potential actions into three *spheres*, namely (i) community manual labor provision, (ii) community financial resource (cash/ in-kind) provision, and (iii) external funding request. The third step was to sum the incidences of all possible actions that the community may have taken within each need *domain* and action *sphere combination*, creating 15 count variables (five domains times three spheres). The fourth step was to apply principal components analysis (PCA) to the resulting data, using nine of the 15 count variables, spanning domain-sphere combinations associated with the infrastructure, health and agriculture domains, which were deemed most relevant in the analysis of the links between cooperative capacity and agricultural productivity.¹³ Hence, the information that is fed into the index is inclusive of all public good needs in the relevant domains. The implicit assumption here is that a generalized index of collective action capacity serves as a good proxy for a wide range of public goods provision.

4.3. Diversity and Similarity Indices

The empirical tests of the model hypotheses require the construction of measures of diversity within the general population and measures of similarity between the leadership and the general population.

We construct four common measures of diversity for the leadership and the general population. The diversity dimensions of choice are in part driven by data availability, and include gender, age, ethnicity, and education. For each dimension, we compute the inverse Simpson concentration index, which captures the diversity in the shares of the dimension-specific groups in a given population. The index is defined as $1/\left[\sum_{g=1}^G (s_g)^2\right]$, where g identifies four demographic dimensions of a group in a given population, G corresponds to the total number of groups in the population, and s_g is the share of the dimension-specific group g in the population. The dimension-specific groups are (i) male and female for gender, (ii) age groups defined at five-year intervals in the range of 15-99, (iii) 14 groups for ethnicity (including “other”), and (iv) no diploma, primary diploma, lower-secondary diploma and secondary diploma for education.

¹³ The results are robust to alternative formulations of the index that differ in terms of the scope of the variables and that are presented in Table A2 of the Appendix. The results obtained with any alternative formulation of the index are available upon request.

While the shares for dimension-specific groups in the general population are based on the sample of all household members captured by the IHS3, the shares for dimension-specific groups within the leadership are informed by the roster of individuals that were interviewed as part of the focus group for the IHS3 Community Questionnaire. A fifth diversity index on landholdings is constructed solely for the general population as the information on the leadership landholdings was not available. Given the continuous nature of the variable, we capture it as difference between the 10th and 90th percentile of the distribution of household landholdings in each community.

The diversity indices for the general population are measures of heterogeneity within the community and should be control variables in the multivariate analysis of collective action capacity. The ethnicity, education and landholding diversity indices for the general population are featured explicitly as independent variables. The gender and age diversity indices for the general population do not enter into the analysis as they are highly correlated across communities.

Finally, the similarity between the leadership and the general population in terms of gender, age, ethnicity and education are captured as the ratios of leadership to community diversity (Jost, 2006). Since we wish to capture the overlap between the leadership and the general population diversity and are not concerned with the comparisons of diversity, the similarity indices are constructed using the greatest diversity index (the leadership or the general population) as the denominator.

4.4. Additional Control Variables

Transactions costs of collective action are hypothesized to be lower where more diverse people actively participate in community organizations. To explore this hypothesis, we rely on the share of female members and share of members under the age of 30 in parent-teachers associations, education committees and health committees that might exist in the community.¹⁴ Additionally, we control for median estimates of education and age for the leadership to account for the potential separate impacts exerted by these variables (through leadership capacity) on collective action capacity. The inclusion of number of households and its squared term as an independent variable accommodates the possibility that very small communities face difficulties in high per

¹⁴ The IHS3 Community Questionnaire solicited information on the presence and membership dynamics of various local organizations, including VDC, agricultural cooperative, farmers' group, savings and credit cooperative, business association, women's group, youth group, political group, religious group, cultural group, health committee, education committee, parent-teacher association, sports group, NGO, community police/watch group, and disabled association.

household contributions for many public goods, whereas higher negotiation, monitoring and enforcement costs can offset initial gains from increasing members from a low base.

The final set of variables captures external conditions that can either favor local goods provision or promote substitution away from these goods. Agro-ecological and climatic variables include agro-ecological zone dummies, and community medians for: slope, annual rainfall, median temperature during the growing season, and the coefficient of variation in rainfall. Community characteristics that may affect relative benefits and costs to public goods provision include (i) categories for the fraction of community members who temporarily migrate out for work, (ii) whether outsiders in-migrate to the community for temporary work, (iii) whether an agricultural extension worker resides in the community, (iv) whether telephone service is available within the community, (v) whether a micro-finance institution is located in the community, and (vi) the distance from the community to the nearest Agricultural and Development Marketing Corporation (ADMARC) outlet.¹⁵

As noted above, we also have access to an administrative data set on supra-community level institutions, which we have matched with the household and community level surveys. At the EPA level, we have information on the number of farmer groups in operating in the area. Such organizations can either complement local collective action by increasing returns to local public goods provision, or they can substitute for collective action failures. At the district level, we have the number of operational donor- and government-funded agricultural development projects and the per household measures of (i) metric tons of subsidized fertilizer during the 2007/08 agricultural season, and (ii) total wages paid as part of MASAF during the 2008/09 agricultural season. As with farmer groups, directly productive inputs, social safety nets, and agricultural project resources may all either substitute or compliment local public goods provision.

Table 2 reports the descriptive statistics. The average values for median community plot maize production per hectare and median community household real annual consumption expenditures per capita stand at 1,118 and 40,774, respectively. In terms of the composition of the collective capacity index, we observe that the transportation and infrastructure domain features the greatest number of potential actions that could have been taken, followed by the education, health, agriculture and law domains. The percentages of communities that actually addressed the need for at least one local public good through collective action in the domains of transportation and infrastructure, education, health, agriculture and law were 41, 41, 27, 5 and 5 percent, respectively. The greatest degree of heterogeneity within the leadership as well as the general population, and the lowest degree of similarity between the leadership and the general population are both in terms of age. The averages for median leader age, share of female leaders and share of female members in the total membership across health committees, education committees and

¹⁵ ADMARC is a parastatal organization mandated to market agricultural produce and inputs. Typically at least one ADMARC outlet could be found in each district center in Malawi.

parent-teacher associations are estimated at 44.2, 27 percent and 43 percent, respectively. An average community is composed of nearly 240 households and approximately 8.5 kilometers away from the nearest ADMARC. While 51 percent of the communities report having received temporary in-migrants for employment, 64 percent also have at least a quarter of their population temporarily migrating out throughout the year.

4.5. Estimation Strategy

The primary and secondary objectives of the analysis are, respectively, (i) to estimate a model of community collective capacity to provide local public goods as a function of variables that can be directly mapped to the theoretical model, and (ii) to estimate the impact of collective capacity on community level agricultural productivity and welfare.

To fulfill the primary objective, an Ordinary Least Squares (OLS) regression of the composite collective capacity index is estimated in which the independent variables correspond to those that were detailed in Sections 4.3 and 4.4 and that extend from the theoretical model.

To realize the secondary objective, the following OLS regression could be estimated:

$$Y_i = \alpha + \beta * CCA_i + \theta * X_i + \varepsilon_i \quad [5]$$

where i denotes community; Y is the outcome variable of interest; CCA is the composite collective capacity index; X is a vector of observable correlates of the outcome of interest; and ε is the error term. This approach assumes that the systematic differences among communities across the continuum of collective capacity are captured through the inclusion of observables.

However, since there is likely unobserved heterogeneity that jointly determines collective action and productivity/welfare outcomes at the community level (i.e. ε in equation [5] is correlated with CCA), the coefficient β in equation [5] is expected to be biased. One solution to this problem is the use of a Two-Stage Least Squares (TSLS) estimation, where the idea is to isolate the movements in CCA that are uncorrelated with ε by finding at least one instrumental variable that predicts collective capacity for local public good provision but exerts no direct impact on the outcomes. This in turn permits a consistent estimation of the β coefficient.

The TSLS approach to involves the estimation of the following linear regressions:

$$CCA_i = \alpha + \gamma * Z_i + \theta * X_i + u_i \quad [6]$$

$$Y_i = \alpha + \beta * CCA_i + \delta * X_i + v_i \quad [7]$$

where i denotes community; Y is either median plot maize production per hectare or median household real annual consumption expenditures per capita; CCA is the endogenous variable of interest; Z identifies the set of instrumental variables that only enter into the first stage estimation (i.e. equation 6); X is a vector of observable correlates that are included in both stages of estimation; and u and v are zero-mean error terms. The predicted values for CCA from equation [6] are entered into equation [7] to recover an estimate for the coefficient β , with the assumption that $E(v_i | Z_i) = 0$. The estimation of equation [6] satisfies also the first objective of the paper.

For TSLS regressions to work, each instrument must satisfy two conditions, namely instrumental relevance and instrumental exogeneity. If an instrument is relevant, then the variation in the instrument is related to the variation in the endogenous variable. The exogeneity requirement is that the instrument must be uncorrelated with the outcome variable, conditional on other covariates, i.e. it must not be correlated with the error term in equation [7]. The fulfillment of both requirements would indicate the validity of a particular instrumental variable candidate.

The three instrumental variables that are used in the analysis include the leadership-community gender and age similarity indices, and the share of female members in the total membership across health committees, education committees and parent-teacher associations in the community. Even though tests for instrumental validity exist and will be discussed in the context of our over-identified model in section 5, instrumental validity relies more on the degree to which theoretical arguments are persuasive in asserting that the instruments are strong predictors of the endogenous variable, and that they affect the outcomes of interest only through their impact on community collective capacity for local public good provision.

Given the conceptual model and the expected relationships between the instruments and the endogenous variable that are discussed in section 4.4, it is straightforward to argue that the instruments fulfill the relevance requirement and are expected to be statistically significant predictors of the collective capacity. While the satisfaction of the exogeneity requirement is more difficult to assert, we appeal to the empirical literature to argue that we are able to control for the major factors associated with both greater representation and voice and with better welfare outcomes. The two major factors discussed in the literature review are education levels and, to a lesser extent, geographic isolation. In our results below, we capture both measures of both absolute education levels of leaders in addition to their education levels relative to the underlying population. We also include measures of distance to major market centers, as well as access to non-locally provided services and infrastructure; all of which capture the degree of isolation of a given community. In section 5.3, we discuss robustness checks that include additional control variables, including additional variables that could capture the degree of openness and innovativeness of a community in addition to education levels and a measure of political connectedness. As our analysis is robust to these checks, we now turn to the discussion of the results from our preferred specification.

5. Results

5.1. First-Stage Estimation Results

Table 3 presents the results from the estimations of equation 6, where columns A and B contain the collective capacity regression results that differ only in terms of the community-level attributes that are posited to ameliorate the negative effects of heterogeneity. Column B presents the results from a more parsimonious specification that serves as the first stage for the TSLS estimates reported in Table 4.¹⁶ The sign, magnitude and statistical significance of the coefficients are comparable across Table 3 columns A and B.

The leader-community gender and age similarity indices take on positive coefficients that are statistically significant at the 1 percent level. This finding is in line with our hypotheses and implies that as the leadership reflects both the gender and age distribution within the community, collective action is higher. While the coefficient for the leader-community ethnic similarity index is not statistically significant, the coefficient associated with the leader-community education similarity index is negative and statistically significant at the 10 percent level even after controlling for the median education level among the leaders, whose coefficient is positive and statistically significant at the 1 percent level, as hypothesized. Since most individuals have limited education, the leadership-community education similarity index is relatively high when leadership is also relatively uneducated. Together with the positive impact of median leadership education level on collective capacity, the results indicate that higher levels of education among the leaders both in absolute terms and relative to the general population are associated with greater collective capacity. The share of female members and members under the age of 30 in the total membership across health committees, education committees and parent-teacher associations in the community both assume positive, statistically significant coefficients. These results are also in line with the predictions from the theoretical model.

The three measures of underlying community heterogeneity, namely ethnic and education diversity – captured by the inverse Simpson concentration indices – and the difference between the 10th and 90th percentile of the land ownership distribution, exert statistically insignificant

¹⁶ The focus on the three instrumental variables in Table 3 is driven by the fact that their utilization as a group yielded the most favorable weak identification F statistics across the TSLS models presented here. The utilization of an expanded set of variables that could ameliorate the negative effects of heterogeneity as instrumental variables, however, yields qualitatively similar first-stage and second-stage results, which are available upon request.

impacts on collective capacity, with the exception of the community education diversity index in column B, where the variable is marginally significant at the 10 percent level. These results are consistent with the hypothesis that representative leadership and broad-based participation in community organizations are effective at ameliorating the negative effects of heterogeneity. Furthermore, collective capacity is found to increase both when communities receive in-migrants looking for work, and with greater shares of members temporarily migrating out. A positive impact from in-migrants is consistent with the hypothesis that more dynamic communities with greater exposure to different experiences gain greater collective capacity. Temporary out-migration may increase knowledge and experience, leading to greater cooperative capacity, but is not consistent with the hypothesis that greater exit options will reduce collective capacity. Interestingly, the coefficients for the number of households and its squared term are not statistically significant, indicating that once we control for divergent interests, the sheer number of households alone does not have a separate impact on the dependent variable.

Moreover, collective capacity is positively associated with favorable rainfall and agro-ecological zone placement. This is in line with the hypothesis that more promising agro-climatic conditions increase returns to local public goods and thus favor greater collective capacity. The measure of rainfall variability also assumes a positive coefficient that is statistically significant at the 10 percent level in column B, indicating that collective capacity is higher when communities face greater climate shocks. Lastly, the attributes that would be positively correlated with agricultural productivity and welfare outcomes, such as the presence of an agricultural extension officer in the community and the district metric tons of subsidized fertilizer per household, take on negative and statistically significant coefficients, indicating that these attributes serve as substitutes to collective capacity for local public good provision. On the other hand, district-level MASAF wage funds per household has a positive coefficient that is statistically significant at the 5 percent level in column B. Thus, while access to non-locally-provided inputs appear to substitute for local public goods provision, the availability of extra-local insurance appears to serve as a complement.

5.2. Second-Stage Estimation Results

We now turn to the discussion of the regression results reported in Table 4. While the estimates are reported from the OLS and TSLS regressions of the two dependent variables, namely logarithmic measures of community median plot maize production per hectare and community median household real annual consumption expenditures per capita, we focus on the TSLS results, presented in columns B and D.¹⁷ The collective capacity for local public good provision

¹⁷ We see that the TSLS coefficient associated with the collective capacity index in Table 4 is always higher than its OLS counterpart, and note that improvements in agricultural productivity and living standards that arise from collective action that spurs local public good provision may be realized through additional alternative mechanisms that are “substitutes” for local public good provision and that may have been omitted from our model; particularly

is positively correlated with our measures of agricultural productivity and welfare. The coefficient of interest is statistically significant at the (i) 5 percent level in the regression of log of community median plot maize production per hectare and (i) 1 percent level in the regression of log of community median household real annual consumption expenditures per capita. The findings are in support of the hypothesis that local public goods can have substantial impacts on agricultural production and productivity, with parallel impacts on living standards.

Other noteworthy results from Table 4 concern the community education diversity and difference between the 10th and 90th percentile of the land ownership distribution. The negative and statistically significant effect of the latter variable on agricultural productivity may capture less dense informal insurance mechanisms and labor- and information-sharing arrangements, which are in turn directly negatively associated with agricultural productivity. The positive TSLS coefficient for the community education diversity in both regressions is likely to reflect the positive correlation of the variable with the overall education level in the community, which is in turn positively associated with the outcomes of interest. The community ethnic diversity has a positive coefficient that is statistically significant at the 5 percent level in column D, suggesting that such diversity has additional impacts on improving consumption outcomes, perhaps through greater density of supra-community networks. Lower temperatures, lower rainfall variability and location in more favorable agro-ecological zones are also similarly associated with better agricultural productivity and welfare outcomes.

Finally, the test results presented in the lower panel of Table 4, under columns B and D, form the econometric foundation for our claims regarding instrumental validity. We first discuss the over-identification test results. The null hypothesis of this test is that the instruments are jointly valid, and that the excluded instruments are correctly excluded from the estimated equation. For instrumental validity, and thereby the inability to reject the null hypothesis of the over-identification test, the test statistic should be statistically insignificant. Instrumental validity is proven in our case, at least empirically, by Sargan statistics that are highly statistically insignificant for both models. Complementing the over-identification tests, we report the result from the orthogonality test of each instrumental variable, which is simply a test of the exogeneity of a subset of instruments whose validity may be subject to scrutiny. The null hypothesis of the

since do not control for household-specific substitution strategies. While these substitutes would theoretically have a positive impact on the outcomes of interest, they would be negative correlated with collective capacity to provide local public goods. This implies that the direction of correlation between the endogenous variable and the error term would be negative and that the positive OLS coefficient associated with the endogenous variable would increase following the use of TSLS. One test of this hypothesis is to exclude two likely substitutes that are currently included in the model, namely district-level 2008/09 rainy season measures of metric tons of fertilizer per household and MASAF wage funds per household, and confirm whether the OLS coefficient associated with the collection action capacity index goes down further. The results, which are available upon request, indicate that the OLS coefficient of interest decreases from 0.031 (statistically significant at the 5 percent level) to 0.022 (not statistically significant) in the regression of log of community median plot maize production per hectare, and similarly drops from 0.017 (statistically significant at the 10 percent level) to 0.016 (not statistically significant) in the regression of log of community median household total real annual consumption per capita.

orthogonality test is that both the unsuspected and suspected instruments are valid. The instruments never fail to pass this test, as indicated by the statistically insignificant orthogonality test statistics. In addition, our instruments take on coefficients that carry anticipated signs and that are statistically significant at the 1 percent level in the first stage estimation, whose results are reported in Table 3, column B. The p-values associated with the under-identification test statistics also signal the fulfillment of the instrumental relevance requirement. The Kleibergen-Paap rk Wald F statistics from the weak identification tests stand at 12.03, which lies in between the Cragg-Donald F statistic critical values (under i.i.d. errors) of 9.08 and 13.91 for 5 percent and 10 percent instrumental variable relative bias, respectively (Stock and Yogo, 2005).

5.3. Robustness Checks

In this section, we consider a number of robustness checks, which are reported in Tables A3, A4 and A5 of the Appendix. First, we note that the leadership-community similarity indices that are in part informed by the focus group composition (which is assumed to reflect the leadership in the community) are closely related to the size of the focus group. Since the timing of the interview throughout the 13-month period of fieldwork, and spatial differences in opportunity cost of participating in the focus group, may have implications for the size and composition of the focus group, we included as additional control variables the interview month fixed effects and the community median ganyu (i.e. casual labor) daily wage rate computed from individual-level reporting as part of the Time Use and Labour Module of the Household Questionnaire. The estimates reported in Tables 3 and 4 are robust to the inclusion of these controls.

A further concern may be raised specifically on the exogeneity of the share of female members in the total membership across health committees, education committees and parent-teacher associations. While we defined the variable in the context of non-directly productive groups that are less likely to influence the dependent variables, it may be closely correlated with the variable defined in terms of more directly productive groups, such as village development committee, tobacco clubs, farmer groups, agricultural cooperatives, business associations and savings and credit cooperatives. The results from the preferred specification are robust to using only the similarity indices as the instrumental variables, and the resulting decline in the weak identification F-statistics is negligible.

Despite the scope of the control variables and the arguments presented in Section 4.5, one might still assert that our instruments may be related to the degree to which a community is progressive, is able to spur innovation, and/or lacks institutional inertia. If true, these unobserved factors would affect the outcomes of interest and the exclusion restriction would be violated. To counteract this potential criticism, we introduce additional independent variables into equations 6 and 7, and test the robustness of our estimates from the preferred specification. One of the controls is the average share of lifetime in community computed across the focus group

respondents which is thought to proxy for the degree of orthodox governance among the leaders of the community. Other control variables rely on the information solicited as part of the IHS3 Community Questionnaire and include the shares of households in the community under the following marital arrangements: (i) matrilineal, neolocal, (ii) matrilineal, matrilineal, (iii) matrilineal, patrilineal, (iv) patrilineal, neolocal and (v) patrilineal, patrilineal. Due to limited observations in categories (i) and (iv), we collapse these arrangements into: i) matrilineal, neolocal/matrilineal, (ii) matrilineal, patrilineal, and (iii) patrilineal, neolocal/patrilineal.

Broadly speaking, patrilineal and particularly patrilineal kinship systems tend to disadvantage women in their intra-household relations, for she is uprooted from her own social group and placed in another on top of the generally patriarchal bias of Malawian society (Takane, 2007). Not only does her husband have control over her in many domains, but she has less ability to appeal to her kin in disagreements, and she has fewer rights over the land that she works.¹⁸ Peters (1997) states that while women exercise authority in the domains of life governed by kinship, male authority is the norm beyond these domains. That is, matrilineal social organization should not be taken to imply a lack of patriarchy, particularly as there are key authoritative roles for men in their own matrilineal lines as “guardians” of their sisters and nieces and nephews (Phiri, 1983). Men are said to ‘own’ the matrilineal while women ‘create’ it (Peters, 1997). A matrilineal system, though, can provide women additional bargaining power in relations with her husband, particularly if the couple resides matrilineally (Reniers, 2003).¹⁹ Hence, the vector of covariates accounting for shares of households under the applicable marital arrangements defined above in terms of lineage and locality captures the relative degree of female empowerment in the community, which may capture more progressive and open leadership at least in terms of women’s representation and voice. As shown in Tables A3 through A5, the results from the preferred specifications are qualitatively similar to those obtained through the introduction of these variables as additional controls.

¹⁸ In any discussion of these traditional arrangements, it must be stressed that these are ideal types, and although at times they are applied rigidly, there is normally scope for great flexibility in their application, particularly when faced with land scarcity (as is the case in much of Malawi today) (Takane, 2007).

¹⁹ Historically, there have been significant pressures on matrilineal systems in Malawi, beginning with Ngoni invasions in the late 19th century (Phiri, 1983). Most scholars seem to think that this patrilineal, warlike group adopted much of the culture of the (mainly) Chewa whom they conquered, although there was a dual process of cultural diffusion (Mtika and Doctor, 2002). Thus the matrilineal Chewa practice both uxorilocal and virilocal marriage, and have less distinct lineages than either the Yao or the Tumbuka (Mtika and Doctor, 2002). Later outsiders continued to put pressure on the matrilineal system, including missionaries, who allocated lands to men, and who preached the moral and economic superiority of man (Phiri, 1983). Both colonial and post-independence state practices also increased the importance of the nuclear family over the matrilineal kin group, including tax that needed to be paid in cash (forcing men into wage labor, often on large plantations or through emigration) and agricultural extension services targeted at men and encouraging cash crops (Peters, 1997). Even apart from large-scale pressures on traditional lineage and land-allocation systems, there is extensive evidence of deviation from the ‘rules’ outlined above. Many of these are due to personal relationships – if a man in a patrilineal group were raised by his mother’s family, he might inherit from them rather than from his patrikin, for example (Takane, 2007). Across Malawi, there is also a trend towards individualistic behavior.

6. Conclusion

Existing empirical literature suggests that economic and socio-cultural heterogeneity can impose significant costs on collective capacity to provide public goods and can exert a direct influence on the distribution of benefits and costs associated with any given public good. For instance, given varying preferences and command over resources, men and women may well obtain different net benefits from the provision of public goods, even if they are part of the same household. In addition to generating different net benefits across members, heterogeneity can increase transactions costs of negotiating and enforcing agreements to provide public goods.

The negative effects of heterogeneity can be mitigated through a composition of local leadership that reflects the underlying heterogeneity in the community population, and similarly through active participation of members with different interests in community life. Both mechanisms could improve the ease with which the distributions of costs and benefits across individuals in the community are assessed, and could facilitate the accommodation of heterogeneous interests, while lowering the costs associated with negotiating, monitoring and enforcing agreements underling collective action for public good provision.

On the whole, the theory suggests that broad civic participation is indicative of viable, well-functioning local institutions that facilitate collective action, and is key to equitable realization of benefits associated with participatory development efforts and sustainability of accrued benefits. This paper conceptualizes in a theoretical model the factors that underlie collective action capacity for public good provision, provides empirical tests of the hypotheses stemming from the theoretical model, and estimates, in a Two-Stage Least Squares framework, the impact of collective action capacity for public good provision on community median estimates of maize yield and household consumption expenditures per capita in rural Malawi.

Our results confirm the importance of having women in leadership roles and in being active members in local organizations. These results are consistent with the hypothesis that women's representation and voice enables communities to negotiate and enforce agreements to provide public goods. The same holds true for having young adults actively participate in organizations and have representation in leadership. However, in Malawi, there is no statistically significant impact of leadership that better reflects ethnic composition of the community. In addition, communities that have more educated leaders – both absolutely and relative to the underlying population –have greater provision of public goods. The results also show that greater public goods provision increases agricultural productivity and consumption per capita. Taken together, these results reinforce the premise that greater representation and participation leads to better outcomes for community members.²⁰

²⁰ Even with (i) the data-intensive nature of our study that brings together a rich set of household survey, geospatial and administrative data, and (ii) the theoretical arguments that are put forth in defense of our instruments and that

The open question that remains is how external interventions and policy mechanisms can foster greater representation and participation in communities where such representation and participation might be limited. As identified in the literature review, there are four broad mechanisms for inducing broad-based representation and participation: (i) legally mandating that certain groups are represented in the leadership and/or the membership of certain projects and organizations at the local level (e.g. the panchayat system in India as discussed by Dreze and Sen, 2002), (ii) recommending increased representation, through a legal, regulatory scheme, (iii) using facilitators to work with communities to increase representation and participation, similar to the strategy commonly employed by community-driven development projects for encouraging inclusive decision making and greater participation of marginalized groups (e.g. the GoBifo project in Sierra Leone, as evaluated by Casey, Glennerster, and Miguel, 2011), and (iv) more broad-based information dissemination programs, e.g. radio programs, that discuss the benefits to inclusive, broad-based development. Externally mandating representation and participation might be sufficient, particularly if enforced extra-locally. However, it may well be that additional actions targeting community norms, and taking into account current vested interests in the status quo, may in fact yield more favorable results more quickly; this is likely to be particularly true in communities where both members and leaders have limited education. Because of the continuing importance of local public goods provision and their potential to significantly improve agricultural productivity and livelihoods, it remains important to gather additional empirical evidence on which specific mechanisms can indeed “induce” the participation of women and other members of diverse groups in local governance.

are supported through the comprehensive set of results from the instrumental variable tests in our toolkit and the robustness checks presented in Tables A3, A4 and A5 in the Appendix, we acknowledge the possibility that the instrumental variables may not adequately satisfy the exogeneity requirement. If there are omitted variables that induce greater openness and acceptance of broad-based participation, that were still not accounted for in robustness checks, that are proxied by the instrumental variables and that would also exert a direct effect on the second stage outcomes, our results could be biased.

References

- Agarwal, B. (2009). "Gender and forest conservation: the impact of women's participation in community forest governance." *Ecological Economics*, 68, pp. 2785-2799.
- Alesina, A. and La Ferrara, E. (2000). "Participation in heterogeneous communities." *Quarterly Journal of Economics*, 115.3, pp. 847-904.
- Alesina, A. and La Ferrara, E. (2005). "Ethnic diversity and economic performance." *Journal of Economic Literature*, 43.3, pp. 762-800.
- Andersson, K., and Agrawal, A. (2011). "Inequalities, institutions, and forest commons." *Global Environmental Change*, 21.3, pp. 866-875.
- Baland, J. M., Bardhan, P., Das, S., and Mookherjee, D. (2010). "Forests to the people: decentralization and forest degradation in the Indian Himalayas." *World Development*, 38.11, pp. 1642-1656.
- Baland, J. M., and Platteau, J. P. (1996). *Halting degradation of natural resources: is there a role for rural communities?* Rome, Italy: Food and Agriculture Organization of the United Nations.
- Baland, J. M., and Platteau, J. P. (1997). "Wealth inequality and efficiency in the commons part I: the unregulated case." *Oxford Economic Papers*, 49.4, pp. 451-482.
- Bardhan, P. (1993). "Symposium on democracy and development." *Journal of Economic Perspectives*, 7.3, pp. 45-49.
- Bardhan, P. (2000). "Irrigation and cooperation: an empirical analysis of 48 irrigation communities in South India." *Economic Development and Cultural Change*, 48.4, pp. 847-865.
- Bardhan, P., and Dayton-Johnson, J. (2002). "Unequal irrigators: heterogeneity and commons management in large-scale multivariate research." *The Drama of the Commons*, pp. 87-112.
- Barham, J., and Chitemi, C. (2009). "Collective action initiatives to improve marketing performance: lessons from farmer groups in Tanzania." *Food Policy*, 34, pp. 53-59.
- Begossi, A. (1998). "Cultural and ecological resilience among Caiaaras of the Atlantic Forest and Caboclos of the Amazon, Brazil." In: Folke, C. and Berkes, F. (eds.), *Linking social and ecological systems*, pp. 129-157. Cambridge, UK: Cambridge University Press.
- Bromley, D., General Editor. Feeny, D., McKean, M., Peters, P., Gilles, J.L., Oakerson, R.J., Runge, C.F., and Thomson, J.T., Coeditors. (1992). *Making the commons work: theory, practice and policy*. San Francisco, CA: Institute for Contemporary Studies Press.

- Casey, K., Glennerster, R., and Miguel, E. (2011). "The GoBifo project evaluation report: assessing the impacts of community driven development in Sierra Leone." *Final report submitted to the World Bank*.
- Collier, P. (2001). "Implications of ethnic diversity." *Economic Policy*, 32, pp. 129-166.
- Chinsinga, B. (2008). "Decentralisation and poverty reduction in Malawi – a critical appraisal." Chapter 3. In: Crawford, G. and Hartmann, C. (eds.), *Decentralisation in Africa: a pathway out of poverty and conflict?* Amsterdam, The Netherlands: Amsterdam University Press.
- Chiweza, A.L. (2010). "A review of the Malawi decentralisation process: lessons from selected districts." Retrieved from:
http://www.concern-universal.org/files/decentralisation_report_molgrd_and_cu_final.pdf.
- Dayton-Johnson, J. (2000). "The determinants of collective action on the local commons: a model with evidence from Mexico." *Journal of Development Economics*, 62.1, pp. 181-208.
- Dayton-Johnson, J., and Bardhan, P. (2002). "Inequality and conservation on the local commons: a theoretical exercise." *Economic Journal*, 112.481, pp. 577-602.
- De Janvry, A., McCarthy, N., and Sadoulet, E. (1998). "Endogenous provision and appropriation in the commons." *American Journal of Agricultural Economics*, 80.3, pp. 658-664.
- Dreze, J., and Sen, A. K. (2002). *India: development and participation*. New York, NY: Oxford University Press.
- Gebremedhin, B., Pender, J., and Tesfay, G. (2004). "Collective action for grazing land management in crop–livestock mixed systems in the highlands of Northern Ethiopia." *Agricultural Systems*, 82.3, pp. 273-290.
- Godquin, M., and Quisumbing, A. "Separate but equal? the gendered nature of social capital in rural Philippine communities." *Journal of International Development*, 20, pp. 13-33.
- Johnson, R. N., and Libecap, G. D. (1982). "Contracting problems and regulation: the case of the fishery." *American Economic Review*, 72.5, pp. 1005-1022.
- Jost, L. (2006). "Entropy and diversity." *Okios*, 113, pp. 363-375.
- Karaivanov, A. (2009). "Heterogeneity, returns to scale and collective action." *Canadian Journal of Economics*, 42.2, pp. 771-807.
- La Ferrara, E. (2002). "Self-help groups and income generation in the informal settlements of Nairobi." *Journal of African Economies*, 11.1, pp. 61-89.

- Mansuri, G., and Rao, V. (2012) "Localizing development: does participation work?" World Bank Policy Research Report. Retrieved from http://siteresources.worldbank.org/INTRES/Resources/469232-1321568702932/8273725-1352313091329/PRR_Localizing_Development_full.pdf
- McCarthy, N. (2004). *With C. Dutilly-Diane, Drabo Boureima, A. Kamara and J.-P. Vanderlinden. Managing resources in erratic environments: an analysis of pastoral systems in Ethiopia, Niger and Burkina Faso.* IFPRI Research Report 135. Washington, DC: International Food Policy Research Institute.
- McCarthy, N., Dutilly-Diané, C., and Drabo, B. (2004). "Cooperation, collective action and natural resources management in Burkina Faso." *Agricultural Systems*, 82.3, pp. 233-255.
- McCarthy, N. and Essam, T. (2009). "Impact of water user association on agricultural productivity in Chile." IFPRI Discussion Paper 892.
- Miguel, E., and Gugerty, M. K. (2005). "Ethnic diversity, social sanctions, and public goods in Kenya." *Journal of Public Economics*, 89.11, pp. 2325-2368.
- Mtika, M., and Doctor, H. (2002). "Matriliny, patriliney and wealth flow variations in rural Malawi." *African Sociological Review*, 6.2, pp. 71-97.
- Naidu, S. C. (2009). "Heterogeneity and collective management: evidence from common forests in Himachal Pradesh, India." *World Development*, 37.3, pp. 676-686.
- Narayan, D., and Pritchett, L. (1999). "Cents and sociability: household income and social capital in rural Tanzania." *Economic Development and Cultural Change*, 47.4, pp. 871-897.
- Olson, M. 1965. *The logic of collective action.* Cambridge, MA: Harvard University Press.
- Ostrom, E. (1990). *Governing the commons: the evolution of institutions for collective action.* Cambridge, UK: The Cambridge University Press.
- Ostrom, E. and Walker, J. (eds). (2003). *Trust and reciprocity: interdisciplinary lessons from experimental research.* New York, NY: Russell Sage Foundation.
- Pandolfelli, L., Meinzen-Dick, R., and Dohrn, S. (2008). "Gender and collective action: motivations, effectiveness and impact." *Journal of International Development*, 20, pp. 1-11.
- Peters, P. (1997). "Against the odds. Matriliny, land and gender in the Shire Highlands of Malawi." *Critique of Anthropology*, 17.2, pp. 189-210.
- Phiri, K. (1983). "Some changes in the matrilineal family system among the Chewa of Malawi since the nineteenth century." *Journal of African History*, 24.2, pp. 257-274.

- Reniers, G. (2003). "Divorce and remarriage in rural Malawi." *Demographic Research*, Special Collection 1, pp. 176-206.
- Ruttan, L. M. (2008). "Economic heterogeneity and the commons: effects on collective action and collective goods provisioning." *World Development*, 36.5, pp. 969-985.
- Sandler, T. (1992). *Collective action: theory and applications*. Ann Arbor, MI: The University of Michigan Press.
- Sen, A. K. (2009). *The idea of justice*. Cambridge, MA: Harvard University Press.
- Stock, J., and Yogo, M. (2005). "Testing for weak instruments in linear IV regression." In: Andrews, D. W. K., and Stock, J. H. (eds.), *Identification and inference for econometric models: essays in honor of Thomas Rothenberg*. Cambridge: Cambridge University Press, pp. 80-108.
- Takane, T. (2007). "Customary land tenure, inheritance rules, and smallholder farmers in Malawi." Institute of Developing Economies Discussion Paper No. 104.
- Varughese, G., and Ostrom, E. (2001). "The contested role of heterogeneity in collective action: some evidence from community forestry in Nepal." *World Development*, 29.5, pp. 747-765.
- Were, E., Roy, J., and Swallow, B. (2008). "Local organisation and gender in water management: a case study from the Kenya highlands." *Journal of International Development*, 20, pp. 69-81.
- Westermann, O., Ashby, J., and Pretty, J. (2005). "Gender and social capital: the importance of gender differences for the maturity and effectiveness of natural resource management groups." *World Development*, 33.11, pp. 1783-1799.

Table 1: Community- & Individual-Level Descriptive Statistics on Focus Group Composition

Variable	Mean	Std. Dev.	Min	Max
<i>Community-Level Statistics (Sample: 628 EAs)</i>				
Size of the Focus Group	8.50	2.48	4	12
Village Chief/Group Village Chief/Traditional Authority: Present in Focus Group †	0.91	0.29	0	1
Village Chief/Group Village Chief/Traditional Authority /Counselor to Village Chief/VDC or ADC Member: Present in Focus Group †	0.98	0.14	0	1
<i>Individual-Level Statistics (Sample: 5,336 Focus Group Respondents)</i>				
Female	0.28	0.45	0	1
Age (Years)	44.95	14.84	15	99
% of Lifetime Spent in Community	0.79	0.32	0	1
<i>Highest Educational Qualification</i>				
None †	0.56	0.50	0	1
Primary †	0.22	0.41	0	1
Junior Secondary †	0.11	0.31	0	1
Secondary & Above †	0.11	0.31	0	1
<i>Ethnicity</i>				
Chewa †	0.55	0.50	0	1
Nyanja †	0.11	0.31	0	1
Yao †	0.09	0.28	0	1
Tumbuka †	0.09	0.29	0	1
Lomwe †	0.04	0.20	0	1
Nkhonde †	0.01	0.08	0	1
Ngoni †	0.03	0.16	0	1
Sena †	0.03	0.17	0	1
Nyakyusa †	0.00	0.05	0	1
Tonga †	0.02	0.15	0	1
Lambya †	0.02	0.13	0	1
Senga †	0.00	0.04	0	1
Sukwa †	0.01	0.09	0	1
Other†	0.00	0.05	0	1
<i>Role in Community</i>				
Village Chief/Group Village Chief/Traditional Authority †	0.22	0.41	0	1
Counselor to Village Chief/Member of ADC VDC †	0.27	0.45	0	1
School Headmaster/Teacher †	0.07	0.25	0	1
Health Worker †	0.03	0.18	0	1
Extension Officer/Lead Farmer/Estate Manager †	0.01	0.11	0	1
Political Leader †	0.00	0.07	0	1
Religious Leader †	0.11	0.31	0	1
Community-Based Organization/Committee Leader/Member †	0.10	0.30	0	1
Business Woman/Man †	0.08	0.28	0	1
Police †	0.02	0.14	0	1
Other †	0.08	0.27	0	1

Note: † identifies a dummy variable.

Table 2: Community-Level Descriptive Statistics

Variable	Mean	Std. Dev.	Min	Max
<i>Outcomes</i>				
Community Median Plot Maize Production (KGs) Per (GPS-Based) Hectare	1,118	559	56	3,918
Log Community Median Plot Maize Production (KGs) Per (GPS-Based) Hectare	6.87	0.61	4.0	8.3
Community Median Household Real Annual Consumption Expenditures Per Capita (Kwacha)	40,774	16,158	10,395	151,951
Log Community Median Household Real Annual Consumption Expenditures Per Capita (Kwacha)	10.55	0.37	9.2	11.9
<i>Collective Action Capacity Index & Underlying Components</i>				
Collective Action Capacity Index	0.00	1.00	-0.7	5.9
<i>Transportation Infrastructure</i>				
Collective Action & Addressing Road/Bridge Need †	0.41	0.49	0.0	1.0
Total # of Outfunding Attempts Across Needs	1.66	3.04	0.0	26.0
Total # of Manpower Provision Instances Across Needs	0.74	0.91	0.0	5.0
Total # of Financial Resource Provision Instances Across Needs	0.29	0.66	0.0	5.0
<i>Agriculture</i>				
Collective Action & Addressing Agricultural Need †	0.05	0.22	0.0	1.0
Total # of Outfunding Attempts Across Needs	0.22	0.84	0.0	10.0
Total # of Manpower Provision Instances Across Needs	0.07	0.27	0.0	2.0
Total # of Financial Resource Provision Instances Across Needs	0.02	0.15	0.0	2.0
<i>Health</i>				
Collective Action & Addressing Health/Water Need †	0.27	0.44	0.0	1.0
Total # of Outfunding Attempts Across Needs	1.48	2.68	0.0	20.0
Total # of Manpower Provision Instances Across Needs	0.49	0.78	0.0	4.0
Total # of Financial Resource Provision Instances Across Needs	0.36	0.68	0.0	4.0
<i>Education</i>				
Collective Action & Addressing School Need †	0.41	0.49	0.0	1.0
Total # of Outfunding Attempts Across Needs	1.61	2.88	0.0	25.0
Total # of Manpower Provision Instances Across Needs	0.66	0.73	0.0	4.0
Total # of Financial Resource Provision Instances Across Needs	0.55	0.72	0.0	4.0
<i>Law</i>				
Collective Action & Addressing Law Need †	0.05	0.21	0.0	1.0
Total # of Outfunding Attempts Across Needs	0.05	0.37	0.0	4.0
Total # of Manpower Provision Instances Across Needs	0.05	0.23	0.0	1.0
Total # of Financial Resource Provision Instances Across Needs	0.02	0.13	0.0	1.0
<i>Leadership Heterogeneity</i>				
Leader Gender Inverse Simpson Concentration Index	1.57	0.35	1.0	2.0
Leader Age Inverse Simpson Concentration Index	4.82	1.29	1.5	9.0
Leader Ethnic Inverse Simpson Concentration Index	1.24	0.43	1.0	4.5
Leader Education Inverse Simpson Concentration Index	2.07	0.72	1.0	3.9
<i>Community Heterogeneity</i>				
Community Gender Inverse Simpson Concentration Index	1.97	0.04	1.7	2.0
Community Age Inverse Simpson Concentration Index	8.33	1.39	3.3	12.0
Community Ethnic Inverse Simpson Concentration Index	0.30	0.23	0.0	0.8
Community Education Inverse Simpson Concentration Index	1.62	0.54	1.0	3.8
Community Difference 10th & 90th Percentile of Land Ownership Distribution	1.08	0.62	0.0	4.9
<i>Accommodating Heterogeneity</i>				
Leadership-Community Gender Similarity Index	0.79	0.17	0.5	1.0
Leadership-Community Age Similarity Index	0.59	0.17	0.2	1.0
Leadership-Community Ethnic Similarity Index	0.74	0.21	0.3	1.0

Table 2 (Cont'd)

Variable	Mean	Std. Dev.	Min	Max
Leadership-Community Education Similarity Index	0.70	0.18	0.3	1.0
Share of Female Members in Health/Education Committee, PTA	0.43	0.14	0.0	1.0
Share of Under 30 Members in Health/Education Committee, PTA	0.20	0.21	0.0	1.0
<i>Additional Leadership Attributes</i>				
Share of Female Leaders	0.27	0.18	0.0	0.9
Median Age Level Among Leaders	44.19	8.61	23.5	74.0
Median Education Level Among Leaders	1.59	0.79	1.0	4.0
<i>Community Attributes Affecting Costs of Negotiating, Monitoring & Enforcing Agreements</i>				
Number of Households in Community (Divided by 10)	23.81	10.69	2.8	82.1
Temporary In-Migration for Employment Opportunities †	0.51	0.50	0.0	1.0
Share of Community Members Temporarily Migrating Out: <1/4 †	0.36	0.48	0.0	1.0
Share of Community Members Temporarily Migrating Out: 1/4 †	0.19	0.39	0.0	1.0
Share of Community Members Temporarily Migrating Out: 1/2 & Above †	0.45	0.50	0.0	1.0
<i>Agro-Ecological, Climatological Attributes Affecting Benefits & Costs of Public Good Provision</i>				
Community Slope (Percent)	4.92	4.33	0.9	29.3
Community Average Temperature of the Wettest Quarter: 1960-1990	235.73	20.20	188	285
Rainy Season (11Nov-31Apr) Precipitation Coefficient of Variation: 1996-2011	0.25	0.04	0.2	0.3
Community Average Annual Rainfall: 2001-2011 (Millimeters)	841.22	93.21	680	1,232
Agro-Ecological Zone: Tropic-warm/semiarid †	0.48	0.50	0.0	1.0
Agro-Ecological Zone: Tropic-warm/subhumid †	0.35	0.48	0.0	1.0
Agro-Ecological Zone: Tropic-cool/semiarid †	0.11	0.31	0.0	1.0
Agro-Ecological Zone: Tropic-cool/subhumid †	0.07	0.25	0.0	1.0
<i>Infrastructure & Services Affecting Benefits & Costs of Public Good Provision</i>				
Telephone in Community †	0.25	0.43	0.0	1.0
Micro-Finance Institution in Community †	0.07	0.26	0.0	1.0
Agricultural Extension Officer in Community †	0.31	0.46	0.0	1.0
Community Distance to Nearest ADMARC (Kilometers)	8.47	5.77	0.5	37.3
# of Farmer Groups in Extension Planning Area	17.49	17.09	0.0	87.0
# of Agriculture Development Projects in District	0.78	1.11	0.0	3.0
District Metric Tons of Subsidized Fertilizer Per Household - 2008/09 Season	0.08	0.03	0.0	0.2
District MASAF Wage Funds Per Household - 2008/09 Season (Divided by 1000)	0.35	0.24	0.1	1.5
Observations	628			

Note: † identifies a dummy variable.

Table 3: OLS Regression Results
Dependent Variable: Collection Action Capacity Index

Variables	A	B_‡
<i>Accommodating Heterogeneity</i>		
Leadership-Community Gender Similarity Index	0.765*** (0.206)	0.794*** (0.205)
Leadership-Community Age Similarity Index	0.714*** (0.188)	0.710*** (0.188)
Leadership-Community Ethnic Similarity Index	0.115 (0.263)	
Leadership-Community Education Similarity Index	-0.396* (0.230)	
Share of Female Members in Health/Education Committee, PTA	0.695** (0.297)	0.952*** (0.266)
Share of Under 30 Members in Health/Education Committee, PTA	0.533** (0.210)	
<i>Community Heterogeneity</i>		
Community Ethnic Inverse Simpson Concentration Index	-0.068 (0.225)	-0.177 (0.157)
Community Education Inverse Simpson Concentration Index	-0.081 (0.078)	-0.136* (0.082)
Community Difference 10th & 90th Percentile of Land Ownership Distribution	0.017 (0.055)	0.008 (0.057)
<i>Additional Leadership Attributes</i>		
Median Age Level Among Leaders	-0.010*** (0.004)	-0.011*** (0.004)
Median Education Level Among Leaders	0.244*** (0.069)	0.284*** (0.067)
<i>Community Attributes Affecting Costs of Negotiating, Monitoring & Enforcing Agreements</i>		
Number of Households in Community (Divided by 10)	-0.013 (0.014)	-0.011 (0.015)
Number of Households in Community (Divided by 10) Squared	0.000 (0.000)	0.000 (0.000)
Temporary In-Migration for Employment Opportunities †	0.420*** (0.090)	0.404*** (0.087)
Share of Community Members Temporarily Migrating Out: 1/4 † ⇨	0.222* (0.116)	0.174 (0.113)
Share of Community Members Temporarily Migrating Out: 1/2 & Above † ⇨	0.178** (0.088)	0.121 (0.084)
<i>Agro-Ecological, Climatological Attributes Affecting Benefits & Costs of Public Good Provision</i>		
Community Slope (Percent)	0.002 (0.007)	0.000 (0.007)
Community Average Temperature of the Wettest Quarter: 1960-1990	-0.001 (0.003)	-0.002 (0.003)
Rainy Season (11Nov-31Apr) Precipitation Coefficient of Variation: 1996-2011	3.438** (1.645)	3.305* (1.708)
Community Average Annual Rainfall: 2001-2011 (Millimeters)	0.001** (0.001)	0.001** (0.001)

Table 3 (Cont'd)

Variables	A	B‡
Agro-Ecological Zone: Tropic-warm/semiarid † ✧	0.110 (0.137)	0.156 (0.139)
Agro-Ecological Zone: Tropic-warm/subhumid † ✧	-0.173 (0.127)	-0.167 (0.131)
Agro-Ecological Zone: Tropic-cool/semiarid † ✧	0.417** (0.186)	0.436** (0.188)
<i>Infrastructure & Services Affecting Benefits & Costs of Public Good Provision</i>		
Telephone in Community †	0.004 (0.099)	0.008 (0.098)
Micro-Finance Institution in Community †	-0.139 (0.165)	-0.147 (0.163)
Agricultural Extension Officer in Community †	-0.161** (0.079)	-0.153* (0.081)
Community Distance to Nearest ADMARC (Kilometers)	0.003 (0.008)	0.001 (0.008)
# of Farmer Groups in Extension Planning Area	-0.002 (0.002)	-0.001 (0.002)
# of Agriculture Development Projects in District	0.087** (0.041)	0.087** (0.042)
District Metric Tons of Subsidized Fertilizer Per Household - 2008/09 Season	-6.477*** (1.695)	-6.334*** (1.634)
District MASAF Wage Funds Per Household - 2008/09 Season (Malawian Kwacha, Divided by 1000)	0.301* (0.155)	0.311** (0.152)
Constant	-2.612** (1.180)	-2.641** (1.205)
Observations		616
R2	0.252	0.238

Note: ‡ indicates the preferred first stage specification for the subsequent TSLS estimations; *** p<0.01, ** p<0.05, * p<0.1; Bootstrapped standard errors reported in parentheses, with 100 replications and districts defined as the strata; † identifies a dummy variable; ⚡ The comparison category is Share of Community Members Temporarily Migrating Out: <1/4; ✧ The comparison category is Agro-Ecological Zone: Tropic-cool/subhumid.

Table 4: OLS & TSLS Regression Results*Columns A & B - Dependent Variable: Log Community Median Plot Maize Production Per Hectare**Columns C & D - Dependent Variable: Log Community Median Household Real Annual Consumption Expenditures Per Capita*

	A - OLS	B - TSLS	C - OLS	D - TSLS
<i>Instrumented Variable</i>				
Collective Action Capacity Index	0.031** (0.016)	0.159** (0.080)	0.017* (0.011)	0.159*** (0.060)
<i>Community Heterogeneity</i>				
Community Ethnic Inverse Simpson Concentration Index	0.046 (0.091)	0.058 (0.096)	0.146** (0.067)	0.159** (0.072)
Community Education Inverse Simpson Concentration Index	0.080** (0.035)	0.091** (0.037)	0.224*** (0.031)	0.237*** (0.034)
Community Difference 10th & 90th Percentile of Land Ownership Distribution	-0.074** (0.032)	-0.071** (0.033)	0.035 (0.022)	0.037 (0.024)
<i>Additional Leadership Attributes</i>				
Median Age Level Among Leaders	-0.006*** (0.002)	-0.004* (0.002)	-0.000 (0.001)	0.001 (0.002)
Median Education Level Among Leaders	0.018 (0.028)	-0.017 (0.035)	0.024 (0.020)	-0.014 (0.025)
<i>Community Attributes Affecting Costs of Negotiating, Monitoring & Enforcing Agreements</i>				
Number of Households in Community (Divided by 10)	0.008 (0.007)	0.010 (0.007)	0.007 (0.005)	0.009 (0.006)
Number of Households in Community (Divided by 10) Squared	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Temporary In-Migration for Employment Opportunities †	0.067* (0.035)	0.009 (0.053)	0.118*** (0.028)	0.054 (0.035)
Share of Community Members Temporarily Migrating Out: 1/4 † ✎	-0.164*** (0.047)	-0.185*** (0.053)	-0.140*** (0.033)	-0.163*** (0.036)
Share of Community Members Temporarily Migrating Out: 1/2 & Above † ✎	-0.196*** (0.040)	-0.213*** (0.039)	-0.080*** (0.024)	-0.098*** (0.030)
<i>Agro-Ecological, Climatological Attributes Affecting Benefits & Costs of Public Good Provision</i>				
Community Slope (Percent)	-0.005 (0.004)	-0.005 (0.004)	0.006** (0.003)	0.007** (0.003)

Table 4 (Cont'd)

	A - OLS	B - TSLS	C - OLS	D - TSLS
Community Average Temperature of the Wettest Quarter: 1960-1990	-0.012*** (0.001)	-0.011*** (0.001)	-0.004*** (0.001)	-0.003*** (0.001)
Rainy Season (11Nov-31Apr) Precipitation Coefficient of Variation: 1996-2011	-3.680*** (0.688)	-4.129*** (0.897)	-2.473*** (0.482)	-2.969*** (0.564)
Community Average Annual Rainfall: 2001-2011 (Millimeters)	0.001*** (0.000)	0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)
Agro-Ecological Zone: Tropic-warm/semiarid † ✧	0.395*** (0.068)	0.360*** (0.074)	0.173** (0.069)	0.134* (0.072)
<i>Agro-Ecological, Climatological Attributes Affecting Benefits & Costs of Public Good Provision</i>				
Agro-Ecological Zone: Tropic-warm/subhumid † ✧	0.333*** (0.058)	0.341*** (0.062)	0.292*** (0.062)	0.302*** (0.061)
Agro-Ecological Zone: Tropic-cool/semiarid † ✧	0.126* (0.068)	0.061 (0.091)	0.109 (0.070)	0.036 (0.087)
<i>Infrastructure & Services Affecting Benefits & Costs of Public Good Provision</i>				
Telephone in Community †	0.062* (0.037)	0.060 (0.040)	0.097*** (0.028)	0.094*** (0.031)
Micro-Finance Institution in Community †	0.164** (0.069)	0.185** (0.074)	0.010 (0.055)	0.032 (0.067)
Agricultural Extension Officer in Community †	-0.004 (0.041)	0.017 (0.042)	0.037 (0.028)	0.061** (0.031)
Community Distance to Nearest ADMARC (Kilometers)	-0.005 (0.004)	-0.005 (0.004)	0.000 (0.002)	0.000 (0.002)
# of Farmer Groups in Extension Planning Area	0.002** (0.001)	0.002* (0.001)	0.002*** (0.001)	0.002*** (0.001)
# of Agriculture Development Projects in District	0.046*** (0.016)	0.036* (0.020)	0.023* (0.013)	0.013 (0.016)
District Metric Tons of Subsidized Fertilizer Per Household - 2008/09 Season	1.513** (0.683)	2.390*** (0.878)	0.290 (0.524)	1.257* (0.732)
District MASAF Wage Funds Per Household - 2008/09 Season (Malawian Kwacha, Divided by 1000)	-0.243*** (0.081)	-0.282*** (0.085)	0.037 (0.048)	-0.006 (0.059)

Table 4 (Cont'd)

	A - OLS	B - TSLS	C - OLS	D - TSLS
Constant	9.356*** (0.542)	9.359*** (0.601)	11.665*** (0.386)	11.668*** (0.456)
Observations		616		
R2	0.544	0.507	0.408	0.288
Instrumental Variables				
Leadership-Community Gender Similarity Index		✓		✓
Leadership-Community Age Similarity Index		✓		✓
Share of Female Members in Health/Education Committee, PTA		✓		✓
Overidentification Test of All Instruments (P-value)		0.660		0.926
Orthogonality Test for Leadership-Community Gender Similarity Index (P-value)		0.481		0.711
<i>Overidentification Test of Instruments Excluding Leadership-Community Gender Similarity Index (P-value)</i>		0.562		0.896
Orthogonality Test for Leadership-Community Age Similarity Index (P-value)		0.386		0.929
<i>Overidentification Test of Instruments Excluding Leadership-Community Age Similarity Index (P-value)</i>		0.778		0.702
Orthogonality Test for Share of Female Members in Health/Education Committee, PTA (P-value)		0.938		0.751
<i>Overidentification Test of Instruments Excluding Share of Female Members in Health/Education Committee, PTA (P-value)</i>		0.363		0.817
Underidentification Test (P-value)		0.000		0.000
Weak Identification Test (Kleibergen-Paap rk Wald F statistic)		12.032		12.032

Note: *** p<0.01, ** p<0.05, * p<0.1; Bootstrapped standard errors reported in parantheses, with 100 replications and districts defined as the strata; † identifies a dummy variable; ♣The comparison category is Share of Community Members Temporarily Migrating Out: <1/4; ♠The comparison category is Agro-Ecological Zone: Tropic-cool/subhumid.

Table A1: Mapping of Variables & Data Sources

Variable	Source
All Variables Listed in Table 1	IHS3 Community Data
Community Median Plot Maize Production (KGs) Per (GPS-Based) Hectare	IHS3 Agricultural Data
Community Median Household Real Annual Consumption Expenditures Per Capita (Kwacha)	IHS3 Household Data
Collective Action Capacity Index & ALL Underlying Components	IHS3 Community Data
Leader Gender Inverse Simpson Concentration Index	IHS3 Community Data
Leader Age Inverse Simpson Concentration Index	IHS3 Community Data
Leader Ethnic Inverse Simpson Concentration Index	IHS3 Community Data
Leader Education Inverse Simpson Concentration Index	IHS3 Community Data
Community Gender Inverse Simpson Concentration Index	IHS3 Household Data
Community Age Inverse Simpson Concentration Index	IHS3 Household Data
Community Ethnic Inverse Simpson Concentration Index	IHS3 Household Data
Community Education Inverse Simpson Concentration Index	IHS3 Household Data
Community Difference 10th & 90th Percentile of Land Ownership Distribution	IHS3 Agricultural Data
Leadership-Community Gender Similarity Index	IHS3 Community & Household Data
Leadership-Community Age Similarity Index	IHS3 Community & Household Data
Leadership-Community Ethnic Similarity Index	IHS3 Community & Household Data
Leadership-Community Education Similarity Index	IHS3 Community & Household Data
Share of Female Members in Health/Edu Comm & PTA	IHS3 Community Data
Share of Under 30 Members in Health/Edu Comm & PTA	IHS3 Community Data
Share of Female Leaders	IHS3 Community Data
Median Age Level Among Leaders	IHS3 Community Data
Median Education Level Among Leaders	IHS3 Community Data
Number of Households in Community (Divided by 10)	IHS3 EA Listing Data
Temporary In-Migration for Employment Opportunities †	IHS3 Community Data
Share of Community Members Temporarily Migrating Out: <1/4 †	IHS3 Community Data
Share of Community Members Temporarily Migrating Out: 1/4 †	IHS3 Community Data
Share of Community Members Temporarily Migrating Out: 1/2 & Above †	IHS3 Community Data
Community Slope (Percent)	IHS3 Household & Geospatial Data
Community Average Temperature of the Wettest Quarter: 1960-1990	IHS3 Household & Geospatial Data
Rainy Season (11Nov-31Apr) Precipitation Coefficient of Variation: 1996-2011	IHS3 Household & Geospatial Data
Community Average Annual Rainfall: 2001-2011 (Millimeters)	IHS3 Household & Geospatial Data
Agro-Ecological Zone: Tropic-warm/semiarid †	IHS3 Household & Geospatial Data
Agro-Ecological Zone: Tropic-warm/subhumid †	IHS3 Household & Geospatial Data
Agro-Ecological Zone: Tropic-cool/semiarid †	IHS3 Household & Geospatial Data
Agro-Ecological Zone: Tropic-cool/subhumid †	IHS3 Household & Geospatial Data
Telephone in Community †	IHS3 Community Data
Micro-Finance Institution in Community †	IHS3 Community Data
Agricultural Extension Officer in Community †	IHS3 Community Data
Community Distance to Nearest ADMARC (Kilometers)	IHS3 Household & Geospatial Data
# of Farmer Groups in Extension Planning Area	Farmers Union of Malawi
# of Agriculture Development Projects in District	Ministry of Agriculture & Food Security
District Metric Tons of Subsidized Fertilizer Per Household - 2008/09 Season	Ministry of Agriculture & Food Security
District MASAF Wage Funds Per Household - 2008/09 Season (Malawian Kwacha, Divided by 1000)	Ministry of Agriculture & Food Security

Note: † identifies a dummy variable.

Table A2: Collective Action Capacity Index Formulations For Sensitivity Analyses

Domain	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
<i>Transportation Infrastructure</i>																
Collective Action & Addressing Road/Bridge Need †									X	X	X	X	X	X	X	X
Total # of Outfunding Attempts Across Needs	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Total # of Manpower Provision Instances Across Needs	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Total # of Financial Resource Provision Instances Across Needs	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Agriculture</i>																
Collective Action & Addressing Agricultural Need †									X	X	X	X	X	X	X	X
Total # of Outfunding Attempts Across Needs	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Total # of Manpower Provision Instances Across Needs	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Total # of Financial Resource Provision Instances Across Needs	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Health</i>																
Collective Action & Addressing Health/Water Need †									X	X	X		X			
Total # of Outfunding Attempts Across Needs	X	X	X	X					X	X	X		X			
Total # of Manpower Provision Instances Across Needs	X	X	X	X					X	X	X		X			
Total # of Financial Resource Provision Instances Across Needs	X	X	X	X					X	X	X		X			
<i>Education</i>																
Collective Action & Addressing School Need †									X	X		X		X		
Total # of Outfunding Attempts Across Needs		X	X		X	X			X	X		X		X		
Total # of Manpower Provision Instances Across Needs		X	X		X	X			X	X		X		X		
Total # of Financial Resource Provision Instances Across Needs		X	X		X	X			X	X		X		X		
<i>Law</i>																
Collective Action & Addressing Law Need †									X		X	X			X	
Total # of Outfunding Attempts Across Needs		X		X	X		X		X		X	X			X	
Total # of Manpower Provision Instances Across Needs		X		X	X		X		X		X	X			X	
Total # of Financial Resource Provision Instances Across Needs		X		X	X		X		X		X	X			X	

Table A2 (Cont'd)

Domain	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
<i>Transportation Infrastructure</i>																
Collective Action & Addressing Road/Bridge Need †									X	X	X	X	X	X	X	X
Total # of Outfunding Attempts Across Needs																
Total # of Manpower Provision Instances Across Needs	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Total # of Financial Resource Provision Instances Across Needs	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Agriculture</i>																
Collective Action & Addressing Agricultural Need †									X	X	X	X	X	X	X	X
Total # of Outfunding Attempts Across Needs																
Total # of Manpower Provision Instances Across Needs	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Total # of Financial Resource Provision Instances Across Needs	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
<i>Health</i>																
Collective Action & Addressing Health/Water Need †									X	X	X		X			
Total # of Outfunding Attempts Across Needs																
Total # of Manpower Provision Instances Across Needs	X	X	X	X					X	X	X		X			
Total # of Financial Resource Provision Instances Across Needs	X	X	X	X					X	X	X		X			
<i>Education</i>																
Collective Action & Addressing School Need †									X	X		X		X		
Total # of Outfunding Attempts Across Needs																
Total # of Manpower Provision Instances Across Needs		X	X		X	X			X	X		X		X		
Total # of Financial Resource Provision Instances Across Needs		X	X		X	X			X	X		X		X		
<i>Law</i>																
Collective Action & Addressing Law Need †									X		X	X			X	
Total # of Outfunding Attempts Across Needs																
Total # of Manpower Provision Instances Across Needs		X		X	X		X		X		X	X			X	
Total # of Financial Resource Provision Instances Across Needs		X		X	X		X		X		X	X			X	

Note: The column highlighted in grey identifies the preferred collective action capacity index formulation.

Table A3: OLS Regression Results
Dependent Variable: Collection Action Capacity Index

Variables	A‡	B	C	D	E	F	G	H
<i>Accommodating Heterogeneity</i>								
Leadership-Community Gender Similarity Index	0.794*** (0.205)	0.834*** (0.204)	0.784*** (0.205)	0.840*** (0.198)	0.786*** (0.218)	0.825*** (0.217)	0.782*** (0.214)	0.836*** (0.207)
Leadership-Community Age Similarity Index	0.710*** (0.188)	0.687*** (0.190)	0.702*** (0.214)	0.673*** (0.211)	0.674*** (0.192)	0.650*** (0.193)	0.676*** (0.213)	0.646*** (0.212)
Share of Female Members in Health/Education Committee, PTA	0.952*** (0.266)		0.999*** (0.261)		0.919*** (0.267)		0.965*** (0.259)	
<i>Community Heterogeneity</i>								
Community Ethnic Inverse Simpson Concentration Index	-0.177 (0.157)	-0.144 (0.161)	-0.228 (0.165)	-0.197 (0.168)	-0.306* (0.172)	-0.278 (0.177)	-0.355** (0.168)	-0.330* (0.172)
Community Education Inverse Simpson Concentration Index	-0.136* (0.082)	-0.126 (0.083)	-0.063 (0.075)	-0.056 (0.076)	-0.141* (0.082)	-0.132 (0.083)	-0.073 (0.075)	-0.066 (0.076)
Community Difference 10th & 90th Percentile of Land Ownership Distribution	0.008 (0.057)	0.015 (0.059)	0.015 (0.062)	0.023 (0.063)	0.019 (0.054)	0.026 (0.056)	0.028 (0.064)	0.037 (0.064)
<i>Additional Leadership Attributes</i>								
Median Age Level Among Leaders	-0.011*** (0.004)	-0.011*** (0.004)	-0.012*** (0.004)	-0.012*** (0.004)	-0.012*** (0.004)	-0.012*** (0.004)	-0.013*** (0.004)	-0.013*** (0.004)
Median Education Level Among Leaders	0.284*** (0.067)	0.297*** (0.066)	0.317*** (0.061)	0.330*** (0.060)	0.273*** (0.066)	0.287*** (0.066)	0.292*** (0.061)	0.305*** (0.061)
<i>Community Attributes Affecting Costs of Negotiating, Monitoring & Enforcing Agreements</i>								
Number of Households in Community (Divided by 10)	-0.011 (0.015)	-0.007 (0.015)	-0.011 (0.014)	-0.008 (0.014)	-0.008 (0.016)	-0.004 (0.016)	-0.008 (0.015)	-0.005 (0.015)
Number of Households in Community (Divided by 10) Squared	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Temporary In-Migration for Employment Opportunities †	0.404*** (0.087)	0.439*** (0.088)	0.430*** (0.078)	0.468*** (0.077)	0.373*** (0.085)	0.407*** (0.086)	0.399*** (0.080)	0.434*** (0.080)

Table A3 (Cont'd)

Variables	A‡	B	C	D	E	F	G	H
Share of Community Members Temporarily Migrating Out: 1/4 † ▨	0.174 (0.113)	0.190* (0.114)	0.148 (0.103)	0.163 (0.102)	0.194 (0.126)	0.211* (0.126)	0.159 (0.108)	0.174 (0.106)
Share of Community Members Temporarily Migrating Out: 1/2 & Above † ▨	0.121 (0.084)	0.131 (0.082)	0.142 (0.089)	0.149* (0.089)	0.155* (0.088)	0.166* (0.085)	0.175* (0.093)	0.183** (0.093)
<i>Agro-Ecological, Climatological Attributes Affecting Benefits & Costs of Public Good Provision</i>								
Community Slope (Percent)	0.000 (0.007)	0.000 (0.007)	-0.001 (0.008)	-0.001 (0.008)	-0.003 (0.007)	-0.003 (0.007)	-0.004 (0.009)	-0.004 (0.009)
Community Average Temperature of the Wettest Quarter: 1960-1990	-0.002 (0.003)	-0.002 (0.003)	-0.003 (0.003)	-0.003 (0.003)	-0.000 (0.004)	-0.001 (0.004)	-0.002 (0.004)	-0.002 (0.004)
Rainy Season (11Nov-31Apr) Precipitation Coefficient of Variation: 1996-2011	3.305* (1.708)	2.810 (1.754)	3.378* (1.733)	2.888* (1.712)	1.666 (1.799)	1.133 (1.834)	1.827 (1.782)	1.289 (1.767)
Community Average Annual Rainfall: 2001-2011 (Millimeters)	0.001** (0.001)	0.001* (0.001)	0.002** (0.001)	0.001** (0.001)	0.001 (0.001)	0.000 (0.001)	0.001 (0.001)	0.001 (0.001)
Agro-Ecological Zone: Tropic-warm/semiarid † ✧	0.156 (0.139)	0.188 (0.134)	0.184 (0.145)	0.224 (0.143)	0.095 (0.186)	0.125 (0.178)	0.141 (0.171)	0.179 (0.169)
Agro-Ecological Zone: Tropic-warm/subhumid † ✧	-0.167 (0.131)	-0.156 (0.130)	-0.152 (0.145)	-0.134 (0.143)	-0.204 (0.168)	-0.195 (0.165)	-0.173 (0.161)	-0.156 (0.159)
Agro-Ecological Zone: Tropic-cool/semiarid † ✧	0.436** (0.188)	0.480** (0.188)	0.422** (0.194)	0.474** (0.196)	0.444** (0.197)	0.485** (0.197)	0.443** (0.201)	0.493** (0.203)
<i>Infrastructure & Services Affecting Benefits & Costs of Public Good Provision</i>								
Telephone in Community †	0.008 (0.098)	0.016 (0.096)	0.031 (0.096)	0.038 (0.095)	0.035 (0.098)	0.044 (0.096)	0.059 (0.095)	0.067 (0.094)
Micro-Finance Institution in Community †	-0.147 (0.163)	-0.162 (0.167)	-0.195 (0.155)	-0.210 (0.160)	-0.133 (0.163)	-0.145 (0.165)	-0.184 (0.153)	-0.196 (0.157)
Agricultural Extension Officer in Community †	-0.153* (0.081)	-0.158** (0.080)	-0.203** (0.083)	-0.208** (0.084)	-0.151* (0.081)	-0.155* (0.080)	-0.201** (0.084)	-0.205** (0.084)
Community Distance to Nearest ADMARC (Kilometers)	0.001 (0.008)	0.002 (0.008)	0.000 (0.007)	0.002 (0.007)	-0.000 (0.008)	0.001 (0.008)	-0.001 (0.007)	0.000 (0.007)

Table A3 (Cont'd)

Variables	A‡	B	C	D	E	F	G	H
# of Farmer Groups in Extension Planning Area	-0.001 (0.002)	-0.001 (0.002)	-0.001 (0.002)	-0.001 (0.002)	-0.001 (0.002)	-0.001 (0.002)	-0.001 (0.002)	-0.001 (0.002)
# of Agriculture Development Projects in District	0.087** (0.042)	0.078* (0.042)	0.082** (0.042)	0.075* (0.042)	0.102*** (0.039)	0.095** (0.040)	0.097** (0.038)	0.092** (0.038)
District Metric Tons of Subsidized Fertilizer Per Household - 2008/09 Season	-6.334*** (1.634)	-6.688*** (1.635)	-6.862*** (1.523)	-7.230*** (1.522)	-6.354*** (1.658)	-6.706*** (1.671)	-6.881*** (1.496)	-7.245*** (1.488)
District MASAF Wage Funds Per Household - 2008/09 Season (Malawian Kwacha, Divided by 1000)	0.311** (0.152)	0.279* (0.152)	0.385*** (0.144)	0.351** (0.149)	0.165 (0.161)	0.128 (0.160)	0.251 (0.153)	0.212 (0.159)
<i>Potential Factors Affecting Focus Group Size & Composition</i>								
Community Median Ganyu Daily Wage Rate (Malawian Kwacha)			-0.001 (0.001)	-0.001 (0.001)			-0.001 (0.001)	-0.001 (0.001)
Month of Interview: April '10 †			-0.204 (0.285)	-0.238 (0.285)			-0.225 (0.281)	-0.259 (0.281)
Month of Interview: May '10 †			-0.541** (0.237)	-0.583** (0.244)			-0.541** (0.239)	-0.581** (0.246)
Month of Interview: June '10 †			-0.544** (0.258)	-0.582** (0.254)			-0.571** (0.263)	-0.610** (0.257)
Month of Interview: July '10 †			-0.780*** (0.245)	-0.796*** (0.241)			-0.812*** (0.251)	-0.827*** (0.248)
Month of Interview: August '10 †			-0.731*** (0.280)	-0.743*** (0.278)			-0.737*** (0.285)	-0.748*** (0.284)
Month of Interview: September '10 †			-0.412 (0.296)	-0.474 (0.293)			-0.448 (0.291)	-0.509* (0.288)
Month of Interview: October '10 †			-0.522** (0.228)	-0.556** (0.233)			-0.499** (0.232)	-0.531** (0.237)
Month of Interview: November '10 †			-0.453* (0.247)	-0.496** (0.244)			-0.437* (0.253)	-0.477* (0.251)
Month of Interview: December '10 †			-0.692*** (0.230)	-0.695*** (0.229)			-0.681*** (0.237)	-0.685*** (0.236)

Table A3 (Cont'd)

Variables	A‡	B	C	D	E	F	G	H
Month of Interview: January '11 †			-0.605** (0.241)	-0.635*** (0.246)			-0.601** (0.247)	-0.630** (0.251)
Month of Interview: February '11 †			-0.115 (0.259)	-0.160 (0.261)			-0.116 (0.260)	-0.160 (0.262)
Month of Interview: March '11 †			-0.413* (0.240)	-0.434* (0.241)			-0.411* (0.246)	-0.433* (0.248)
<i>Additional Controls</i>								
% of Households Under Matrilineal, Neolocal/Matrilineal Marriage ♦					0.389*** (0.131)	0.404*** (0.132)	0.364*** (0.131)	0.381*** (0.132)
% of Households Under Patrilineal, Neolocal/Matrilineal/Patrilineal Marriage ♦					0.365 (0.286)	0.383 (0.290)	0.384 (0.268)	0.407 (0.269)
Average Leader Share of Lifetime in Community					0.080 (0.237)	0.115 (0.240)	-0.035 (0.299)	-0.011 (0.300)
Constant	-2.641** (1.205)	-1.919 (1.218)	-1.962 (1.349)	-1.207 (1.278)	-2.143 (1.567)	-1.452 (1.575)	-1.263 (1.551)	-0.514 (1.469)
Observations	616	616	610	610	616	616	610	610
R2	0.238	0.222	0.284	0.267	0.250	0.236	0.296	0.280

Note:‡ indicates the preferred first stage specification reported in Table 4; *** p<0.01, ** p<0.05, * p<0.1; Bootstrapped standard errors reported in parantheses, with 100 replications and districts defined as the strata; † identifies a dummy variable; ⌘ The comparison category is Share of Community Members Temporarily Migrating Out: <1/4; ✧ The comparison category is Agro-Ecological Zone: Tropic-cool/subhumid; Month of interview fixed effects include a dummy variable for each month from April 2010 to March 2011, with the dummy variable for March 2010 identifying the comparison group; ♦ The omitted variable is the share of community household population under matrilineal, patrilineal marital arrangement; Average leader share of lifetime spent in community is calculated based on the information that is available for each focus group respondent regarding his/her age and number of years of residence in the community.

Table A4: TSLS Regression Results
Dependent Variable: Log Community Median Plot Maize Production Per Hectare

	A‡	B	C	D	E	F	G	H
<i>Instrumented Variable</i>								
Collective Action Capacity Index	0.159** (0.080)	0.163* (0.096)	0.152* (0.078)	0.181* (0.106)	0.150* (0.082)	0.153 (0.101)	0.141* (0.080)	0.168 (0.107)
<i>Community Heterogeneity</i>								
Community Ethnic Inverse Simpson Concentration Index	0.058 (0.096)	0.058 (0.098)	0.049 (0.101)	0.053 (0.102)	0.031 (0.099)	0.032 (0.102)	0.017 (0.100)	0.025 (0.102)
Community Education Inverse Simpson Concentration Index	0.091** (0.037)	0.092** (0.038)	0.114*** (0.036)	0.115*** (0.037)	0.090** (0.038)	0.090** (0.039)	0.115*** (0.036)	0.116*** (0.037)
Community Difference 10th & 90th Percentile of Land Ownership Distribution	-0.071** (0.033)	-0.071** (0.033)	-0.055* (0.033)	-0.055* (0.033)	-0.069** (0.034)	-0.069** (0.034)	-0.051 (0.033)	-0.052 (0.033)
<i>Additional Leadership Attributes</i>								
Median Age Level Among Leaders	-0.004* (0.002)	-0.004* (0.002)	-0.005** (0.002)	-0.004* (0.003)	-0.004* (0.002)	-0.004* (0.002)	-0.005** (0.002)	-0.005* (0.003)
Median Education Level Among Leaders	-0.017 (0.035)	-0.018 (0.039)	-0.018 (0.038)	-0.027 (0.042)	-0.019 (0.039)	-0.019 (0.042)	-0.017 (0.038)	-0.025 (0.042)
<i>Community Attributes Affecting Costs of Negotiating, Monitoring & Enforcing Agreements</i>								
Number of Households in Community (Divided by 10)	0.010 (0.007)	0.010 (0.007)	0.012* (0.007)	0.012* (0.007)	0.010 (0.007)	0.010 (0.007)	0.012* (0.007)	0.012* (0.007)
Number of Households in Community (Divided by 10) Squared	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Temporary In-Migration for Employment Opportunities †	0.009 (0.053)	0.007 (0.058)	0.017 (0.056)	0.003 (0.065)	0.006 (0.051)	0.005 (0.057)	0.015 (0.055)	0.003 (0.064)
Share of Community Members Temporarily Migrating Out: 1/4 † ▫	-0.185*** (0.053)	-0.186*** (0.054)	-0.194*** (0.043)	-0.198*** (0.047)	-0.177*** (0.055)	-0.178*** (0.057)	-0.182*** (0.044)	-0.187*** (0.048)
Share of Community Members Temporarily Migrating Out: 1/2 & Above † ▫	-0.213*** (0.039)	-0.213*** (0.040)	-0.216*** (0.044)	-0.220*** (0.048)	-0.205*** (0.039)	-0.205*** (0.041)	-0.207*** (0.045)	-0.212*** (0.049)

Table A4 (Cont'd)

	A‡	B	C	D	E	F	G	H
<i>Agro-Ecological, Climatological Attributes Affecting Benefits & Costs of Public Good Provision</i>								
Community Slope (Percent)	-0.005 (0.004)	-0.005 (0.004)	-0.005 (0.005)	-0.005 (0.005)	-0.005 (0.004)	-0.005 (0.004)	-0.005 (0.005)	-0.005 (0.005)
Community Average								
Temperature of the Wettest Quarter: 1960-1990	-0.011*** (0.001)	-0.011*** (0.001)	-0.012*** (0.002)	-0.012*** (0.002)	-0.011*** (0.001)	-0.011*** (0.001)	-0.011*** (0.002)	-0.011*** (0.002)
Rainy Season (11Nov-31Apr)								
Precipitation Coefficient of Variation: 1996-2011	-4.129*** (0.897)	-4.144*** (0.903)	-3.740*** (0.631)	-3.842*** (0.699)	-4.421*** (0.907)	-4.425*** (0.914)	-4.069*** (0.675)	-4.112*** (0.699)
Community Average								
Annual Rainfall: 2001-2011 (Millimeters)	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)	0.001** (0.000)	0.001*** (0.000)	0.001** (0.000)
Agro-Ecological Zone: Tropic-warm/semiarid † ✧	0.360*** (0.074)	0.359*** (0.073)	0.419*** (0.076)	0.410*** (0.084)	0.343*** (0.077)	0.343*** (0.076)	0.391*** (0.079)	0.386*** (0.085)
Agro-Ecological Zone: Tropic-warm/subhumid † ✧	0.341*** (0.062)	0.341*** (0.062)	0.394*** (0.075)	0.396*** (0.078)	0.328*** (0.067)	0.328*** (0.066)	0.372*** (0.072)	0.375*** (0.075)
Agro-Ecological Zone: Tropic-cool/semiarid † ✧	0.061 (0.091)	0.058 (0.097)	0.099 (0.102)	0.084 (0.114)	0.063 (0.092)	0.062 (0.099)	0.098 (0.105)	0.085 (0.116)
<i>Infrastructure & Services Affecting Benefits & Costs of Public Good Provision</i>								
Telephone in Community †	0.060 (0.040)	0.060 (0.041)	0.055 (0.043)	0.054 (0.045)	0.063 (0.040)	0.063 (0.040)	0.058 (0.042)	0.056 (0.045)
Micro-Finance Institution in Community †	0.185** (0.074)	0.186** (0.075)	0.175*** (0.066)	0.181*** (0.069)	0.185** (0.076)	0.185** (0.077)	0.173*** (0.066)	0.178*** (0.069)
Agricultural Extension Officer in Community †	0.017 (0.042)	0.017 (0.043)	0.008 (0.041)	0.014 (0.044)	0.016 (0.041)	0.016 (0.042)	0.007 (0.041)	0.012 (0.044)
Community Distance to Nearest ADMARC (Kilometers)	-0.005 (0.004)	-0.005 (0.004)	-0.004 (0.003)	-0.004 (0.003)	-0.006 (0.004)	-0.006 (0.004)	-0.005 (0.003)	-0.004 (0.003)
# of Farmer Groups in Extension Planning Area	0.002* (0.001)	0.002* (0.001)	0.002** (0.001)	0.002* (0.001)	0.002** (0.001)	0.002* (0.001)	0.002** (0.001)	0.002** (0.001)
# of Agriculture Development Projects in District	0.036* (0.020)	0.036* (0.021)	0.048*** (0.015)	0.046*** (0.017)	0.038* (0.021)	0.038* (0.022)	0.050*** (0.016)	0.047*** (0.018)

Table A4 (Cont'd)

	A‡	B	C	D	E	F	G	H
District Metric Tons of Subsidized Fertilizer Per Household - 2008/09 Season	2.390*** (0.878)	2.419*** (0.935)	2.129** (0.949)	2.347** (1.163)	2.364*** (0.890)	2.380** (0.955)	2.106** (0.960)	2.303** (1.169)
District MASAF Wage Funds Per Household - 2008/09 Season (Malawian Kwacha, Divided by 1000)	-0.282*** (0.085)	-0.283*** (0.088)	-0.270*** (0.084)	-0.281*** (0.092)	-0.306*** (0.085)	-0.306*** (0.086)	-0.297*** (0.082)	-0.303*** (0.087)
Constant	9.359*** (0.601)	9.359*** (0.606)	9.546*** (0.598)	9.527*** (0.617)	9.390*** (0.669)	9.390*** (0.683)	9.488*** (0.676)	9.463*** (0.689)
Community Median Ganyu Daily Wage Rate & Month of Interview Fixed Effects Included	NO	NO	YES	YES	NO	NO	YES	YES
Average Leader Share of Lifetime Spent in Community & Shares of Community Household Population Under Each Marital Regime Included	NO	NO	NO	NO	YES	YES	YES	YES
Observations	616	616	610	610	616	616	610	610
R2	0.507	0.505	0.543	0.524	0.513	0.512	0.551	0.534
Instrumental Variables								
Leadership-Community Gender Similarity Index	✓	✓	✓	✓	✓	✓	✓	✓
Leadership-Community Age Similarity Index	✓	✓	✓	✓	✓	✓	✓	✓
Share of Female Members in Health/Education Committee, PTA	✓		✓		✓		✓	
Overidentification Test of All Instruments (P-value)	0.660	0.365	0.837	0.780	0.667	0.369	0.862	0.787
Underidentification Test	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Weak Identification Test (Kleibergen-Paap rk Wald F statistic)	12.032	11.808	12.415	11.740	11.248	11.020	11.760	11.167

Note:‡ indicates the TSLS estimates reported in Table 5; *** p<0.01, ** p<0.05, * p<0.1; Bootstrapped standard errors reported in parentheses, with 100 replications and districts defined as the strata; † identifies a dummy variable; ♠ The comparison category is Share of Community Members Temporarily Migrating Out: <1/4; ♡ The comparison category is Agro-Ecological Zone: Tropic-cool/subhumid.

Table A5: TSLS Regression Results
Dependent Variable: Log Community Median Household Real Annual Consumption Expenditures Per Capita

	A‡	B	C	D	E	F	G	H
<i>Instrumented Variable</i>								
Collective Action Capacity Index	0.159*** (0.060)	0.146* (0.075)	0.174*** (0.050)	0.182*** (0.068)	0.134** (0.060)	0.104 (0.075)	0.144*** (0.049)	0.135** (0.066)
<i>Community Heterogeneity</i>								
Community Ethnic Inverse Simpson Concentration Index	0.159** (0.072)	0.158** (0.072)	0.139* (0.072)	0.140* (0.075)	0.104 (0.072)	0.097 (0.071)	0.084 (0.070)	0.081 (0.074)
Community Education Inverse Simpson Concentration Index	0.237*** (0.034)	0.235*** (0.034)	0.230*** (0.034)	0.230*** (0.035)	0.242*** (0.032)	0.239*** (0.032)	0.240*** (0.032)	0.240*** (0.032)
Community Difference 10th & 90th Percentile of Land Ownership Distribution	0.037 (0.024)	0.037 (0.024)	0.011 (0.027)	0.011 (0.028)	0.045* (0.024)	0.045* (0.023)	0.020 (0.026)	0.020 (0.026)
<i>Additional Leadership Attributes</i>								
Median Age Level Among Leaders	0.001 (0.002)	0.001 (0.002)	0.001 (0.002)	0.002 (0.002)	0.001 (0.002)	0.001 (0.002)	0.002 (0.002)	0.002 (0.002)
Median Education Level Among Leaders	-0.014 (0.025)	-0.010 (0.026)	-0.042 (0.027)	-0.044 (0.031)	-0.019 (0.026)	-0.011 (0.028)	-0.034 (0.027)	-0.031 (0.030)
<i>Community Attributes Affecting Costs of Negotiating, Monitoring & Enforcing Agreements</i>								
Number of Households in Community (Divided by 10)	0.009 (0.006)	0.009 (0.006)	0.006 (0.006)	0.006 (0.006)	0.008 (0.005)	0.008 (0.005)	0.005 (0.006)	0.005 (0.006)
Number of Households in Community (Divided by 10) Squared	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Temporary In-Migration for Employment Opportunities †	0.054 (0.035)	0.060 (0.041)	0.059* (0.035)	0.055 (0.042)	0.049 (0.033)	0.061 (0.038)	0.059* (0.033)	0.063 (0.040)
Share of Community Members Temporarily Migrating Out: 1/4 † ▯	-0.163*** (0.036)	-0.161*** (0.035)	-0.110*** (0.038)	-0.111*** (0.039)	-0.126*** (0.037)	-0.120*** (0.036)	-0.068* (0.037)	-0.066* (0.037)
Share of Community Members Temporarily Migrating Out: 1/2 & Above † ▯	-0.098*** (0.030)	-0.096*** (0.030)	-0.102*** (0.028)	-0.103*** (0.029)	-0.083*** (0.029)	-0.078*** (0.029)	-0.086*** (0.028)	-0.084*** (0.028)

Table A5 (Cont'd)

	A†	B	C	D	E	F	G	H
<i>Agro-Ecological, Climatological Attributes Affecting Benefits & Costs of Public Good Provision</i>								
Community Slope (Percent)	0.007** (0.003)	0.007** (0.003)	0.009*** (0.003)	0.009*** (0.003)	0.008*** (0.003)	0.008*** (0.003)	0.009*** (0.002)	0.009*** (0.002)
Community Average Temperature of the Wettest Quarter: 1960-1990	-0.003*** (0.001)	-0.003*** (0.001)	-0.002* (0.001)	-0.002* (0.001)	-0.001 (0.001)	-0.001 (0.001)	-0.000 (0.001)	-0.000 (0.001)
Rainy Season (11Nov-31Apr) Precipitation Coefficient of Variation: 1996-2011	-2.969*** (0.564)	-2.923*** (0.600)	-2.865*** (0.630)	-2.892*** (0.664)	-3.543*** (0.541)	-3.497*** (0.541)	-3.377*** (0.614)	-3.361*** (0.615)
Community Average Annual Rainfall: 2001-2011 (Millimeters)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Agro-Ecological Zone: Tropic-warm/semiarid † ✧	0.134* (0.072)	0.137* (0.073)	0.108 (0.075)	0.105 (0.078)	0.052 (0.072)	0.057 (0.073)	0.015 (0.075)	0.017 (0.076)
Agro-Ecological Zone: Tropic-warm/subhumid † ✧	0.302*** (0.061)	0.301*** (0.062)	0.269*** (0.063)	0.269*** (0.064)	0.236*** (0.063)	0.232*** (0.065)	0.192*** (0.063)	0.191*** (0.064)
Agro-Ecological Zone: Tropic-cool/semiarid † ✧	0.036 (0.087)	0.043 (0.092)	0.047 (0.082)	0.043 (0.092)	0.024 (0.084)	0.038 (0.090)	0.029 (0.078)	0.034 (0.086)
<i>Infrastructure & Services Affecting Benefits & Costs of Public Good Provision</i>								
Telephone in Community †	0.094*** (0.031)	0.095*** (0.031)	0.083*** (0.028)	0.083*** (0.029)	0.089*** (0.030)	0.090*** (0.030)	0.077*** (0.028)	0.077*** (0.029)
Micro-Finance Institution in Community †	0.032 (0.067)	0.030 (0.067)	0.019 (0.057)	0.020 (0.061)	0.016 (0.062)	0.011 (0.061)	0.002 (0.051)	0.000 (0.054)
Agricultural Extension Officer in Community †	0.061** (0.031)	0.059* (0.030)	0.045 (0.031)	0.047 (0.033)	0.060* (0.031)	0.055* (0.030)	0.042 (0.030)	0.040 (0.031)
Community Distance to Nearest ADMARC (Kilometers)	0.000 (0.002)	0.000 (0.002)	0.001 (0.002)	0.001 (0.003)	-0.000 (0.002)	-0.000 (0.002)	0.001 (0.002)	0.001 (0.002)
# of Farmer Groups in Extension Planning Area	0.002*** (0.001)	0.002*** (0.001)	0.002*** (0.001)	0.002*** (0.001)	0.002*** (0.001)	0.002*** (0.001)	0.002*** (0.001)	0.002*** (0.001)
# of Agriculture Development Projects in District	0.013 (0.016)	0.014 (0.017)	0.012 (0.017)	0.011 (0.018)	0.003 (0.016)	0.006 (0.017)	0.004 (0.016)	0.004 (0.016)

Table A5 (Cont'd)

	A‡	B	C	D	E	F	G	H
District Metric Tons of Subsidized Fertilizer Per Household - 2008/09 Season	1.257* (0.732)	1.169 (0.828)	1.706** (0.711)	1.764** (0.781)	1.351* (0.749)	1.150 (0.837)	1.755** (0.682)	1.683** (0.733)
District MASAF Wage Funds Per Household - 2008/09 Season (Malawian Kwacha, Divided by 1000)	-0.006 (0.059)	-0.002 (0.060)	-0.035 (0.053)	-0.038 (0.058)	-0.039 (0.054)	-0.035 (0.053)	-0.069 (0.050)	-0.066 (0.052)
Constant	11.668*** (0.456)	11.668*** (0.476)	11.661*** (0.415)	11.656*** (0.416)	11.311*** (0.478)	11.312*** (0.494)	11.087*** (0.466)	11.096*** (0.463)
Community Median Ganyu Daily Wage Rate & Month of Interview Fixed Effects Included	NO	NO	YES	YES	NO	NO	YES	YES
Average Leader Share of Lifetime Spent in Community & Shares of Community Household Population Under Each Marital Regime Included	NO	NO	NO	NO	YES	YES	YES	YES
Observations	616	616	610	610	616	616	610	610
R2	0.288	0.309	0.371	0.357	0.352	0.390	0.444	0.457
Instrumental Variables								
Leadership-Community Gender Similarity Index	✓	✓	✓	✓	✓	✓	✓	✓
Leadership-Community Age Similarity Index	✓	✓	✓	✓	✓	✓	✓	✓
Share of Female Members in Health/Education Committee, PTA	✓		✓		✓		✓	
Overidentification Test of All Instruments (P-value)	0.926	0.814	0.933	0.755	0.717	0.710	0.934	0.786
Underidentification Test	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Weak Identification Test (Kleibergen-Paap rk Wald F statistic)	12.032	11.808	12.415	11.740	11.248	11.020	11.760	11.167

Note:‡ indicates the TSLS estimates reported in Table 5; *** p<0.01, ** p<0.05, * p<0.1; Bootstrapped standard errors reported in parentheses, with 100 replications and districts defined as the strata; † identifies a dummy variable; ⚡ The comparison category is Share of Community Members Temporarily Migrating Out: <1/4; ⚡ The comparison category is Agro-Ecological Zone: Tropic-cool/subhumid.