Impact Evaluation (IE) Concept Note
The Impact of Targeting Mechanisms on Efficiency and Equity of Irrigation in Mozambique

Mozambique

P154869

August 2015

Keywords: O12, O13, D61, D70, Q18, I38, Q25

Table of Contents

IE PROFILE INDICATORS

No.	Indicator	Description
1	IE code	P154869
2	IE Title	DIME Agriculture Mozambique Irrigation Incentives for Collective Action
3	IE TTL	Florence Kondylis
4	IE Contact Person	Astrid Zwager (DECIE)
5	Region	AFR
6	Sector Board/Global Practice	Water
7	WBG PID	N/A
8	WBG Project Name	N/A
9	Project TTL	N/A
10	Intervention	Small Scale Irrigation systems
11	Main Outcomes	Land holdings, Efficiency of schemes, yield
12	IE Unit of Intervention/Randomization	Village
13	Number of IE Units of Intervention	56 villages
14	IE Unit of Analysis	Plot
15	Number of IE Units of Analysis	1400 farmers
16	Number of Treatment Arms	1
17	IE Question 1 (Treatment Arm 1)	1. Does decentralized targeting achieve a higher average increase in yields from irrigation investments than rule-based targeting? 2. Does rule based targeting lead to a more equitable distribution of benefits and greater inclusion of small landholders than decentralized targeting?
18	Method IE Question 1	Random assignment at the community level
19	Mechanism tested in IE Question 1	Targeting, package
25	Gender-specific treatment (Yes, No)	No
27	Gender analysis (Yes, No)	Yes
28	IE Team & Affiliations	Florence Kondylis (Senior Economist, DECIE) Paul Christian (Research Associate, Cornell University) Teevrat Garg (Research Officer, LSE; Ass. Professor UCSD) Astrid Zwager (IE coordinator, DECIE) Steven Glover (IE Field Coordinator, DECIE)
29	Estimated Budget (including research time)	Total in USD1,1 million
30	CN Review Date	Sept - 2015
31	Estimated Timeframe for IE	September-2015 to September-2018
32	Main Local Counterpart Institution(s)	Ministry of Agriculture, National Irrigation Institute

1. EXECUTIVE SUMMARY

Irrigation is more important than ever in ensuring sustainable livelihoods for farmers In the face of increasing climate uncertainty. Yet, in the poorest parts of the rural world, the rates of irrigation are low. In Mozambique, only 8% of all farmers have access to irrigation. One source of Mozambique's inability to develop widespread irrigation infrastructure has been an inability to sustain and protect previous investments in the sector. A lack of attention to creating local institutions capable of maintaining the infrastructure has historically led to infrastructure decaying before reaching its productive potential (FAO, 2005). Forming strong institutions begins with selecting the right set of beneficiaries who can function and cooperate together to maximize productive potential of schemes. The poorest farmers have the lowest rates of access to irrigation (Mather, Cunguara, and Boughton, 2008), meaning that future investments also need to prioritize the involvement of these households.

We propose to study the impact of the placement of irrigation kits in 56 villages covering five to ten hectares of land each in the southern province of Gaza in Mozambique. To uncover the efficacy of different targeting regimes, we will randomize the model of kit deliveries – in half the villages, we will use scorecards to target benefits to small landholders when determining which plots/farmers get the kit, and in the other half we will allow for decentralized targeting by letting the communities determine the placement of the kit in their respective villages.

Using a combination of experimental and quasi-experimental techniques, we will answer three primary questions. First, which targeting mechanism generates the highest average yields? Second, which targeting mechanism achieves better distributional equity and better efficiency in providing kits to small holders? Third, we attempt to uncover how targeting regimes interact perform in the face of positive spillovers in the good being targeted, in this case the irrigation kits.

To understand the impact on yields we will exploit the randomized nature of targeting regime to uncover the average treatment effect of targeting regime. The context further provides us with a unique opportunity to identify the effect of small scale irrigation on yields. We can exploit the fact that random assignment into the scorecard based targeting treatment creates exogenous variation in irrigation status among those who meet the eligibility criteria. By comparing outcomes of people who are assigned to participate in the irrigation in the scorecard based regime with those who would have been eligible through the scorecard but assigned to the community consensus treatment, we can identify the effect of the irrigation on outcomes. To unpack the effect of irrigation on yields, we will exploit the sharp discontinuity in the rules-based targeting by comparing plots just above and just below the threshold.

2. Background and Key Institutional Features

Mozambique has experienced significant growth over the previous decade, with an average annual real GDP growth rate of over 7% (WB). Despite this strong economic performance, poverty rates have remained stubbornly high, especially in rural areas. The most recent National Poverty Assessment in 2010 estimated that around 55% out of almost 21.5m people lived below the poverty line, with no indication that it fell in rural areas since the previous survey in 2002. Recent lack of poverty reduction has been attributed to a failure to increase yields for smallholder farmers throughout the agricultural sector (Jones and Tarp, 2012; Arndt et al., 2012).

Over three-quarters of people in Mozambique depend on agriculture as their primary livelihood (FAO, 2015). The sector is dominated by smallholders (<10 ha), which represent over 99% of production units (MINAG, 2012) and account for 95% of the country's production (FAO, 2015). Average farm size is around 1.4 ha, yet only 10% of the 36 million hectares of arable land is currently being cultivated (MINAG, 2010). Labor shortages and low levels of mechanization are constraints to extensive agricultural growth. Agricultural productivity is extremely low - maize and rice yields are around 0.75 and 0.28 tn/ha respectively (MINAG, 2012), many times below regional and international levels. This is exemplified by the low utilization of agricultural inputs by farmers, where only 9% of farmers used a variety of improved seed, 3% used fertilizer, 6% used pesticides, 7% used animal traction and almost no mechanization.

The majority of agricultural production is rain fed, as such rainfall patterns across agricultural seasons is still the most important factor affecting production variability (Cunguara and Kelly, 2009). This is in particular the case for the central part of the country, where agriculture is dominated by smallholder rain-fed farms which have experienced significant harvest losses as a result from extreme weather like droughts and floods. Mozambique ranks third amongst the African countries most exposed to risks from multiple weather-related hazards, suffering from periodic floods, cyclones and droughts. Drought affects by far the largest number of people and climate change is expected to increase exposure to extreme weather.

Promoting sustainable irrigation and drainage is essential in making smallholder farmers resilient to these intensifying climate variations, and the government of Mozambique has made the development of irrigation a priority for agriculture growth and rural development. With abundant water resources, Mozambique has enormous capacity for irrigated agriculture, with an estimated 3.3 million hectares potentially irrigable (MINAG, 2010). At the start of the 1980s around 120,000 ha was irrigated (MINAG, 2013), yet areas were abandoned in the civil war and many schemes fell into disrepair. Today, rehabilitation efforts have ensured that 50,000 ha are in use. Currently only 8% of farmers have access to any type of irrigation method.

The AfDB Sustainable Land & Water Resource Management Project (SLWRMP) aims to increase the resilience of communities towards the adverse impacts of climate change through the development of 56 small scale irrigation kits, covering between 5 and 10 ha, in the Southern province of Gaza. The Gaza province is one of the most adversely affected provinces in terms of climate change events with frequent occurrence of droughts in the northern parts and floods in the coastal areas of the province. The Project will be implemented in the four drought affected districts of Guija Mabalane, Chicualacuala and Massengena.

Preliminary missions with the project revealed that the process of selecting beneficiaries was a unique challenge for the implementation of this project. Large scale irrigation projects are generally constrained by geographical conditions in where they can be located, but the small scale kits used by the SLWRMP can be flexibly located. Once a village is selected, there will be a range of "eligible area" in which the kit could be feasibly placed. Since the kits cannot cover the entire eligible area, location within this zone must be based the project's goals and priorities for who should benefit and how the kits should be used. Historically, investments in irrigation have left out the poorest households (Mather, Cunguara, and Boughton, 2008) and have not been used effectively because of a lack of investment in institutions to operate and maintain the equipment (FAO, 2005). This experience has led to a debate within the project about what are the most important goals in selecting beneficiaries. On the one hand, ensuring access for the poor is a major priority of the project, which favors using eligibility tests to prioritize beneficiaries from vulnerable households. On the other hand a centralized process risks losing buy-in from communities and will result in the benefits accruing to those less capable of using the equipment and creating institutions, ultimately threatening the efficiency and sustainability of the scheme. The latter being a rational for favoring a community-based decentralized process for selecting beneficiaries.

Irrigation interventions hold tremendous potential to help farmers cope with increasing climate variability, and to ensure food security in many poorer regions of the world. Yet, failure to properly manage irrigation schemes often limits returns to investment in this sector. Increasingly, governance models of irrigation are decentralized to local water users' groups. The idea is that users of a scheme should be more successful in monitoring and organizing the Operation & Maintenance (O&M) activities, as they have the local information and incentives to do so. Yet, the potential for collective action failures remain, and little is known as to which governance arrangement and incentive designs can overcome commons problems.

The proposed evaluation aims to shed light on if different approaches to select beneficiaries can reduce elite capture and result in a more equitable distribution of benefits. However, the potential trade-offs between inclusion of smaller farmers and successful management remain an empirical question. We will exploit exogenous variations in the composition of water users groups induced by the random assignment to the two targeting regimes to shed light on the causal relation between group composition and collective action over operation and maintenance of the schemes, and final impact on production.

The evaluation is part of a wider, long-term research agenda aimed at understanding (1) how to best leverage the irrigation investments to increase resilience of small farmers and (2) how to build local institutions to ensure sustainability of these schemes. The portolfio includes projects in Mozambique, Nepal, Kenya and Rwanda. The research teams are working closer together to identify potential solutions to overcome problems associated with irrigation while adapting them to local context. The other Mozambique evaluation will develop monitoring and accountability interventions to induce higher levels of collective action towards a cooperative, more efficient equilibrium. The evaluation (P154165) is embedded in the WB PROIRRI Sustainable Irrigation Project (P107598).

3. LITERATURE REVIEW

The literature on targeting recognizes that there is generally a tradeoff between the degree to which benefits can be channeled to the most vulnerable and the direct and indirect costs of guaranteeing this kind of targeting. Policymakers may agree that they prefer that limited resources be spent on those who fall below a certain wealth or poverty threshold, but direct costs of ensuring that only these households participant can be high, because identifying these households can be time-consuming and costly (Karlan and Thuysbaert, 2013; Atalas et al, 2013). In addition, targeting rules impose indirect costs if subsidies accrue to those who are not the most efficient users of the technology, limiting the benefits of the technology (Jack, 2013; Basurto, Dupas, and Robinson, 2015). This literature recognizes three constraints on the implementation of efficient and effective targeting: 1.Identifying priority beneficiaries can be time intensive and inaccurate; 2. Project staff do not have the information needed to determine who will be the most productive users of the technology; and 3. If soliciting community input requires giving discretion to the community, the project does not know the degree to which local elites may be willing and able to take advantage of this discretion in order to “capture” benefits for themselves and their connected neighbors even when they are not the most efficient users or the most prioritized type of households.

The literature has made considerable progress in understanding the nature and efficiency of different targeting regimes, shedding light on how to most cost-effectively identify target populations. For instance in an seminal randomized control trial over 640 villages in Indonesia, Alatas et al. (2012) show that community targeting, where the villagers rank everyone from richest to poorest, does worse at identifying the poor than proxy means test (PMT), even though community targeting results in higher overall satisfaction than PMT. These results demonstrate that PMT tests can be an affordable and feasible mechanism for selecting beneficiaries. Using the same data, Atalas et al (2013) show that elite capture is not prevalent or substantial enough to meaningfully affect distribution of benefits. Karlan and Thuysbaert (2013) show that PMT and participatory targeting methods identify similar groups of beneficiaries. Other studies show that contexts with less competition for elite influence can lead to more substantial diversion of benefits to elites that that skews allocations towards those with influence and way from those deemed deserving by policymakers' objective metrics (Acemoglu, Reed, Robinson, 2013). These contributions demonstrate that both PMT and decentralized methods can be cost-effective ways of reaching the poor, but that the risk of elite capture can depend on the context. Recent targeting investigations have explored indirect methods of targeting with low direct costs to the project such as ordeal mechanisms (Nichols, Smolensky and Tideman, 1971; Nichols and Zeckhauser, 1982; Ravallion, 1991; Besley and Coate, 1992, Alatas et al., 2013, Dupas et al 2013) or auctioning benefits to users to reveal willingness to pay (Jack, 2013). Evidence of the success of such mechanisms is mixed. While these mechanisms can induce self-selection and revelation of poor users with minimal costs to the project, they also impose direct costs on the beneficiaries themselves, who must endure the ordeal or pay a fee to satisfy the mechanism.

In addition to building on the literature for using PMT scores or community knowledge to identify poor beneficiaries, our study will contribute to the literature on targeting mechanisms in two ways. First, it expands the current targeting literature from social programs that provide a fixed, certain individual benefit that is both rival and excludable–a pure private good such as a food subsidy—to a productive resource (irrigation infrastructure). Second, irrigation infrastructure also has the features of a public good, meaning that we will be able to trace out the implications of beneficiary selection on collective group behavior. Previous literature has focused on individual, excludable goods like food subsidies (Basurto, Dupas and Robinson, 2015), health subsidies (Dupas, Hoffman, Kremer, and Zwane, 2013), and cash transfers (Atalas et al, 2012). However, many social programs transfer goods that are inputs in productive processes and must be shared as a common pool. In such a context, targeting has ambiguous welfare implications.

Our proposed study builds on the literature in two ways. First, we will provide a theoretical exposition of how targeting productive public goods can alter the efficiency and distributional consequences of community level targeting. Second, we will test this model in the area of irrigation using a field experiment in Mozambique where the determination of placement of irrigation infrastructure was randomized across villages subjected to score-based targeting and villages where communities could decide on their own where the infrastructure was placed. To the best of our knowledge, no research has yet examined this critical question that underlines the consequences of different targeting regimes in the presence of positive externalities.

The fact that irrigation equipment is a productive resource means that targeting experience three types of errors. The first two—including those who are "not poor enough" to be included or excluding those who were poor and should have been included--are well studied in the literature on targeting consumption goods. The third--transferring benefits to those who have a low return and excluding those with a high return--has been less well studied. If the poorest farmers are less efficient users of irrigation equipment than relatively richer farmers, prioritizing delivery to the poorest may lead to a smaller improvement in overall village production with high redistributive costs. Jack (2013) finds that when productive inputs are allocated based on beneficiaries bidding on the inputs at auction the resulting matching of inputs to productive potential leads to productivity gains of around 30% relative to a random distribution of inputs, suggesting that the costs of mis-allocation can be substantial. Allowing the community the option of allocating to relatively wealthier farmers, who may be able to extract larger gains from the technology, may lead to larger overall production gains but at the cost of deleterious redistributive implications.

Basurto, Dupas, and Robinson (2015) highlight differences in decentralized targeting patterns between an agricultural input subsidy and a food subsidy in Malawi. They show that when community elites allocate nutrition subsidies, they more closely approximate the results of hypothetical proxy means tests than when they allocate agricultural input subsidies. They take these results to mean that the elites have information about who would use the agricultural inputs most productively. Given that villages in Malawi share resources extensively, this implies that unlike subsidies of private goods like foods, prioritizing the poor for agricultural input subsidies may not maximize benefits to the poor who would be better off ceding their share of the subsidy to a more productive farmer who may share the resulting bounty. However, Basurto, Dupas, and Robinson do not assign groups to participate in PMT targeting or decentralized targeting, meaning they are not able to observe the outcomes of each or measure the hypothetical efficiency losses associated with PMT targeting or decentralized targeting. Therefore a key parameter for the SLWRMP project is how much productivity is lost by using a PMT to direct subsidies to the poor? How valuable is the information on productivity that local elites have? If the project surrenders some decision making to local elites, to what degree do the elites weight productivity gains against households need and against their own potential private benefits? The findings of Basurto, Dupas, and Robinson suggest that the answers to all these questions determine whether a PMT based mechanism creates greater benefits for the poor than a community-driven, decentralized allocation mechanism, and these questions cannot be answered without a trial.

If policymakers determine that local elites are better able to determine the most productive users of technology than PMTs, they still have to worry about whether the elites will act in the public interest, because elite capture in the placement of public goods is a common phenomenon. Our reading of the literature suggests however that it is possible to limit the risk of capture enough to make using local information an attractive strategy. For instance, Keefer & Khemani (2009) use data in India to show that even though politicians tend to favor "pork-barrel" projects for their constituencies, political constraints can limit or enhance such preferences. Hoffman et al. (2015) show, using experimental data, that local leaders in Kenya favor the ability to target public goods substantially more than control over the allocation of maintenance budgets, highlighting that maintaining political control is prioritized over personal financial benefits. Elite capture in programs that give elites discretion in allocation, therefore, may be directed more towards reduced efficiency than leakage. Thus we complement this growing body of work by understanding the efficiency gains (or losses) from targeting impure public goods.

Little is known about the potential for elite capture and the resulting efficiency of targeting mechanisms when the targeted good generates positive externalities and has some elements of a public good (non-rival and/or non-excludable). Irrigation kits require investments in operations (for example fuel to run the pumps) and maintenance (to repair problems). Since the kit can cover land owned by multiple households who cannot be excluded once the kit is placed and function, there are incentives to free-ride on the investments of others. At the stage of selecting whose land will be covered, irrigation is a common pool resources, since the benefits of using irrigation are not excludable to anyone in the scheme, but assigning more land in the scheme to one person must reduce the coverage of another household's land. Once the kit is installed, operations and maintenance become public goods, where everyone benefits proportionate to their share of land included in the scheme, but costs of contributions must be borne by individuals. The non-excludability of functioning irrigation systems can cause under-provision of the public good (Ostrom, 2003). This result of sub-optimal provision of non-rival goods has led economists to search for contexts that exacerbate this type of collective action problem and ways in which institutional structures can shift the equilibrium group behavior toward the Pareto optimum.

Lizzeri & Persico (2001) show that if providing public goods is a way for local elites to build public support, these goods are underprovided when the elites do not have the ability to target who benefits from the public good, suggesting that allowing some degree of elite capture may actually be welfare maximizing if attention and effort of elites is necessary to ensure collective cooperation. This result suggests that if the PMT method excludes the elites, the functioning of the scheme may suffer because elites are not invested in insuring the functioning of the scheme. Theoretical and experimental investigations of collective action in the provision of public goods and common pool resources have posited that the degree of failure to achieve Pareto optimal outcomes may depend on the group structure. For example, the size of the groups is ambiguously related to the level of contributions toward a public good. Large groups may be better at providing public goods if splitting up the costs is a way of "sharing the burden" or diversifying risk, or small groups may be more effective if they are easier to coordinate or they increase the likelihood that one member is willing to bear the fixed costs of public good provision alone because the personal benefits are sufficiently high. (Olson,1965; Issac, Walker, and Williams, 1994). Using PMTs are likely to increase size of the group and include more poor individuals, since requiring the smallest landholders to be covered will likely increase their representation in these groups. The effect of including them on provision of public goods is ambiguous.

Our reading of the literature suggests that if the project was only interested in maximizing the proportion of households with small plots of land who participate in the subsidy, the cheapest and least risky way to do so without imposing an undue burden on the poor would be to use proxy means tests. However, the literature also suggests that ceding control over allocations to communities may have advantages in this context of subsidizing a productive input that are missed in much of the literature on PMTs which focus on subsidies of private goods. Communities may have information about how to distribute the irrigation equipment in ways that maximize benefits beyond what is achievable by those prioritized by the PMT. The only way for the project to know how great these gains might be and how much they might be undone by elite capture is to try both methods to assess the relative advantages of each.

4. POLICY RELEVANCE

Mozambique is currently participating in the Comprehensive African Agriculture Development Program (CAADP), building upon the groundwork laid down by the Strategic Plan for Agricultural Development 2010-19 (PEDSA). A key component of this agrarian development strategy is a strong focus on sustainable land management and irrigation, Pillar 1 of the CAADP Compact and Investment Plan. To this end, the Ministry of Agriculture has been undertaking several significant irrigation construction and rehabilitation projects with partner organizations. Some of these projects include PROIRRI with the World Bank, SLWRMP and BLIRCP with the African Development Bank, and PROSUL with IFAD.

The Xai Xai targeting project has enormous potential to inform public policy. First, at the local level, the research findings will inform the project and other practitioners on the benefits (and costs) of different targeting regimes when considering the provision of irrigation infrastructure to local communities. Different entities including AfDB are considering expanding this program to other communities (outside the initial target area). The results from this evaluation will allow government and donors to make an informed decision on the use of community-driven or score-based targeting in the placement of these new projects.

Second, we will provide insights and contrast the provision of pure private goods versus common pool resources, and whether the potential for elite capture, leakage and efficiency losses is lower or higher. This could inform not just the current project in the current site, but a variety of programs across the development sphere. For instance, should projects provide individual level private goods or should they provide goods to a community that may have different benefits for different individuals, but could be collectively more efficient?

Third, in the area of irrigation, we can address a long-standing puzzle on how best to provide irrigation amenities to small-scale farmers in the absence of large infrastructure projects. Most previous investigation of the cost-effectiveness of irrigation investments has focused on large scale irrigation projects such as dams and canals (Duflo and Pande, 2007). Our results will provide some of the first rigorous estimates (to our knowledge) of the impact of providing small-scale kits that can be positioned with relative flexibility in terms of geographic location. By comparing irrigation users with non-users and exploiting differences in how beneficiaries are selected, we will be provide cost-benefit estimates on the returns to investment in small-scale irrigation that are much more rigorously grounded than those currently used by the government of Mozambique and others (see Section 11). Or results on formation and functioning of user groups will further demonstrate how to maximize these benefits. These estimates will be able to better direct investment in future irrigation schemes toward maximizing return.

5. DESCRIPTION OF INTERVENTION

The project will provide 56 small scale irrigation kits in the southern province of Gaza. The project will be implemented in the four drought affected districts of Guija Mabalane, Chicualacuala and Massengena. The irrigation systems will consist of sprinkler systems covering either 5 – 10 hectare, depending on the availability of continuous stretches of suitable land in the selected communities. The final list of communities to receive the kits is currently being finalized by the project. The final selection is mainly based on communities being suitable for irrigation. Some initially selected communities turned out to already have irrigation infrastructure in place and were dropped. Due to the selection process, we find it unlikely that non-selected villages are comparable to the final list of selected sites. The minimal plot size per farmer covered by the scheme is 0.5 hectare. As a consequence a maximum of 10 or 20 farmers can benefit from the scheme, depending on its size. Since the number of farmers that are interested will likely exceed the number that can be covered, a within village level selection procedure is necessary. The procedure to determine the exact location and beneficiaries within a village will be done in two steps:

1. Identification of potential area: the project's engineers determine and map the total potential area that can be irrigated in the community. This assessment will take into consideration the technical constraints of the kits (e.g. maximum distance from water source is 500 meters) and suitability for agriculture. Usually this area is larger than the 5 – 10 hectare, the maximum area that can be covered by the kit.
2. Identification of beneficiaries: once the eligible area is determined, villages will follow strict protocols to either select the exact plots to be covered or the farmers that will populate the kit.

The evaluation will test two different approaches to carry out step 2 to understand how different protocols can result in a more equitable distribution of benefits and if there are potential trade-offs between inclusion of smaller farmers and productivity gains. The main features of the models are described below. The detailed draft protocols are currently being piloted and can be provided upon request.

Model A: Score-based targeting

Selection under Model A happens according to a fixed set of criteria for placing the schemes and is designed to prioritize the smaller farmers in the community. Each farmer will irrigate between 0.5 and 1 hectare WITHIN the kit. In addition, the eligibility criteria seek to prioritize farmers that have a TOTAL of up to 2 ha of cultivated land. Based on data from the TIA, a nationally representative sample of rural income generation, this cutoff includes the 72nd percentile of famers in the province where the project operates. Small farmers will be identified using a land proxy test. To construct the score card used to determine the land proxy, we followed the method presented in Shreiner and Lory (2013), which describes the development of a poverty scorecard for Mozambique. We used data from the 2012 National Agriculture Survey (TIA) for the Gaza province. See Annex A for a description of the methods used to construct the score card. The score card can be found in Annex B. The selected area should maximize the number of priority farmers included.

Model B: Decentralized Community-driven targeting

Model B will give the community freedom to decide who they want to benefit from the kit, as well as the proportion of irrigated land each of the beneficiaries will farm. The selection process will be mediated by the project staff and village leaders. The community will be provided with subjective criteria to apply such as:

- Which farmers who own land in the zone would be most committed to maintaining and using the kit?
- Which farmers are practicing or are most interested in practicing market oriented farming practices?
- Are any existing associations with organic structure or individual farmers willing to turn create a legal association to form a water users group?

Ethical issues

The project team considered either of the selection procedures during the design of the project. Ethical issues would arise if we would have strong evidence that either of these models would provide more desirable results. A priori we have no reason to believe that either of the two procedures will result in a better outcome in terms of poverty reduction, welfare gains and/or potential for conflict within the community. In particular we expect to inflict no harm on the participating communities and farmers. We will discuss the study setup with the Cornell University IRB and ask them to review it if deemed necessary.

6. Theory of Change

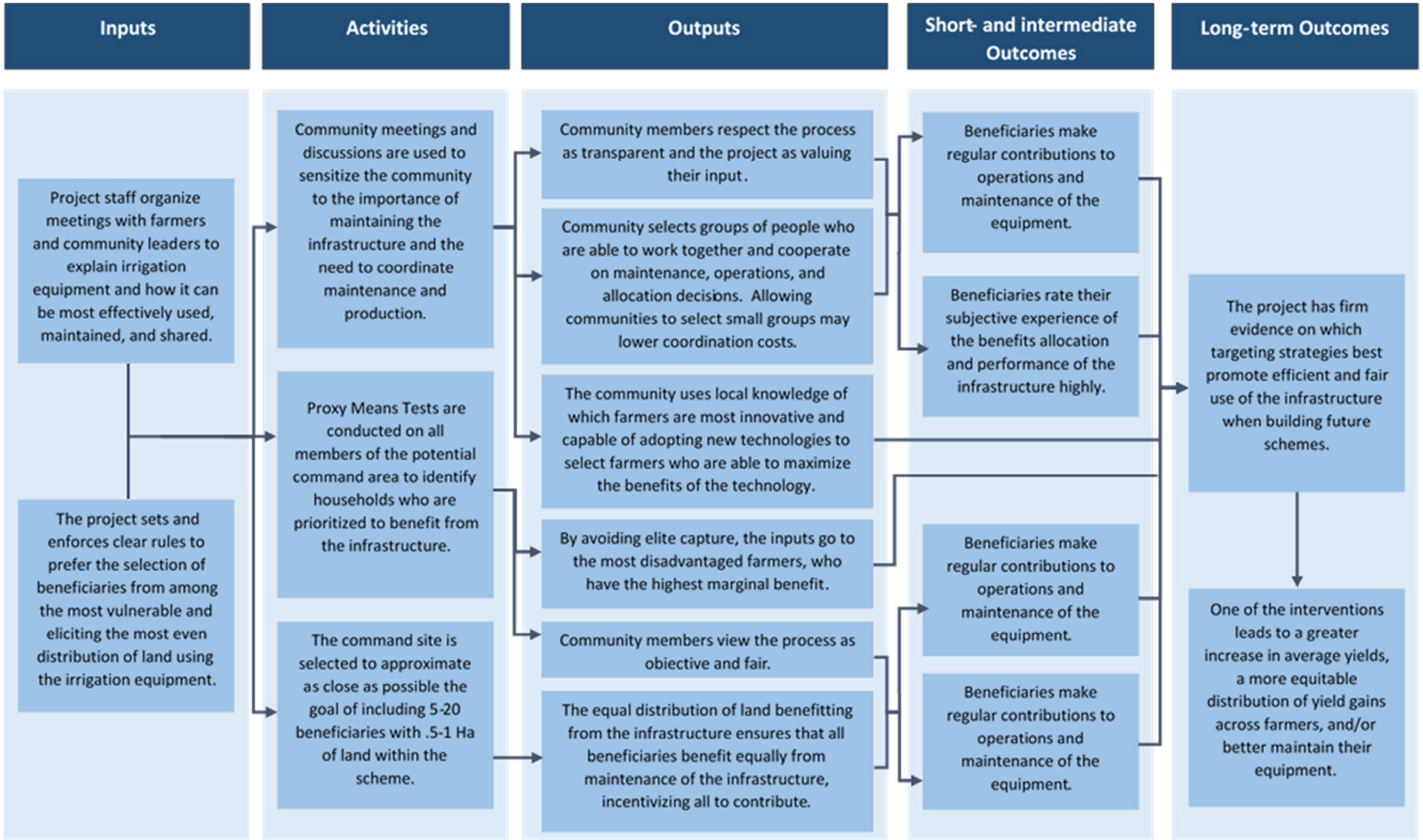
The evaluation is based on comparing the processes by which two possible systems for allocating the infrastructure to potential beneficiaries translates to efficient use of the irrigation equipment, equitable distribution of the benefits, and sustainable use and maintenance of the components. The processes by which these two systems translate to outcomes rely on a set of assumptions whose accuracy will determine which system results in a more effective scheme.

The ideal underlying the decentralized community-driven mechanism is that communities have information about the characteristics of their members that the project would have a difficult time measuring, but which are crucial to the success of the scheme. In particular, communities know which farmers are most innovative and able to apply new farming technologies and so may be able to understand which beneficiaries would use the equipment most productively. They may also know which farmers are most able to work together to form well-functioning users associations and cooperatives that are capable of effectively using the equipment. Finally, engaging the community leaders and trusting their judgment may foster engagement in the process and lead to the community feeling ownership. The crucial assumptions underlying the potential success of this treatment arm relate to the power structure of the village. If bargaining power is skewed in favor of elites or wealthy households, these members may capture the benefits of the system for themselves even if others could create greater value from the technology. The ideal behind this treatment arm assumes that efficiency gains will arise by one of the following three things being true: 1. Village leaders are benevolent enough to consider the total welfare gains of their community when selecting beneficiaries, even if the gains do not directly accrue to them; 2. Village leaders are able to redistribute gains effectively enough *ex* post that they will allocate the equipment to those who benefit most, knowing that they can be compensated for not taking it themselves; 3. Village leaders are the most productive potential users of the technology and so giving it to them maximizes efficiency.

The ideal underlying the rule-based targeting system is that objective and enforced rules are necessary to achieve fair outcomes that favor the most vulnerable. The process envisions that using proxy tests and expressed preferences for equal distributions of benefits will achieve a set of beneficiaries that is generally poorer and an allocation of benefits that is fairer than would likely be achieved without clear rules. The objective, public nature of the test may lead the community to view the process as more transparent and fair than consensus meetings that are subject to elite capture. The assumptions underlying these ideals are that the rules accurately identify priority households, that the rules are enforced by the project, and that the rules are viewed with legitimacy by the community. These assumptions are most likely to hold when the proxy means tests are accurate and well-grounded so that the populations they identify are ones that the communities would also view as vulnerable and worthy of targeting. Furthermore, the success of this system depends on whether trying to achieve a set of beneficiaries with a relatively equal distribution of land will not damage cooperation. It may be the case that having one or a few large landholders involved makes it easier for the scheme to function by creating a natural leader with the incentives to keep the equipment functioning smoothly.

This evaluation is about putting these two theories of change in competition with each other. We anticipate that the decentralized choice may yield higher returns on average in terms of income improvements and sustainability of the scheme but that the rule-based targeting will more effectively concentrate gains among the poor. To assess these claims, we will measure the land-distribution of the eligible area and of the included farmers and baseline agricultural productivity among farmers who could be included in the scheme, monitor the selection of beneficiaries, monitor maintenance and operations of the scheme, and measure productivity gains following the installation of the scheme.

FIGURE 1 – THEORY OF CHANGE



7. HYPOTHESES/EVALUATION QUESTIONS

The proposed study design will seek to answer two main research questions:

Evaluation Question 1: Does decentralized targeting achieve a higher average increase in yields from irrigation investments than rule-based targeting?

> 1.a. Do communities and community leaders choose to give the irrigation benefits to relatively large landholders or relatively small ones?
> 1.b. Do communities allocate irrigation equipment to people who have the highest baseline agricultural productivity?
> 1.c. Do people with higher average baseline productivity experience bigger yield and income improvements from technology than farmers who start at a lower level?

Evaluation Question 2: Does rule based targeting lead to a more equitable distribution of benefits and greater inclusion of small landholders than decentralized targeting?

> 2.a. How well do proxy means tests correctly identify small landholders?
> 2.b. If proxy means tests are imperfect, can elites leverage the imperfections of the tests to capture benefits despite the rules?
> 2.c. Does enforcing an equal distribution of benefits lead to greater coordination costs that threaten the sustainable functioning of the schemes?

The evaluation questions were derived by careful discussion with the project staff about how they planned to design their targeting strategy. Initial discussions revealed a very clear debate between those staff who felt that the project should prioritize the mission of targeting the most vulnerable populations and those who felt that having the project tell the community who could and could not benefit would create frictions that would undermine the success of the project. The staff in favor of decentralizing the process felt that community participation and buy-in from the leaders was the only way to ensure that the equipment would be used correctly and that functioning users associations would be formed. The staff in favor of the rules felt strongly that without clear rules about who could benefit, the elite farmers in the village would capture the entire benefit of the scheme, leaving out the poor members the project was designed to help.

The interventions based on community meetings vs. proxy means test were designed through detailed discussions with the staff and field testing which options were most feasible for achieving each of the two ideals these competing groups of staff had in mind. We hypothesize that the decentralized targeting will lead to smaller groups of beneficiaries who tend to be holders of larger plots of cultivated land and who have higher baseline agricultural productivity relative to a representative sample of the population of the pool of potential beneficiaries. In contrast, we expect that the rule based targeting will lead to a larger average group of beneficiaries who predominately come from the set of smaller landholders from the potential pool and that the largest landholders will not be included. We further expect that the rule-based treatment will achieve a more equal distribution of the land covered by the irrigation equipment across beneficiaries.

The question of which treatment will lead to more sustainable maintenance and operation of the scheme is an open empirical question. On the one hand, theory of collective action says that including one large beneficiary in the group may lead this person to bear the full cost of operating the equipment which would suggest that decentralized treatment may lead to better functioning of the group. On the other hand, equal distribution of benefits may reduce the risk of free-riding which might suggest that the rule-based system would be more effective. So the hypothesize for which type of group will have the most regular and effective contributions to operations and maintenance is ambiguous.

8. MAIN OUTCOMES OF INTEREST

TABLE 1 – MAIN OUTCOMES OF INTEREST

Outcome Type	Outcome Name	Definition	Measurement Level
Primary			
	Yield	Total Revenue per hectare (mean and variance)	Individual/plot
	Profit	Total income received from crop sales net input costs (mean and variance)	Individual/plot
Secondary (maintenance)			
	Efficiency of water pump	Liters/second pumped	Scheme
	Efficiency of scheme	Liters/second provided at the extremity of the scheme	Scheme
	Contributions to O&M	Yearly fee payments	Individual
	Irrigation usage	Number of hours the system is used per day	Scheme
Secondary (targeting)			
	Plot size	Plot size (hectare) within the irrigation kit	Individual
	Equity of plot size	Z score of plot size (hectare) within the irrigation kit	Individual
	Number of beneficiaries	Number of farmers covered by the kit	Scheme
	Total cultivated land	Sum of area of all cultivated plots of beneficiaries (hectare)	Individual
	Within scheme land distribution	Z score of total cultivated land of beneficiaries	Scheme
	Land distribution	Z score of total cultivated land in the eligible area	Community
	Proxy means score	Result from land proxy questionnaire	Individual
	Satisfaction	Perceived fairness of the selection procedure	Individual

9. Evaluation Design and Sampling Strategy

9.1 Treatment and Control Groups

The evaluation will aim to respond the research questions by testing two different approaches to select beneficiaries within the selected community. The identification strategy will be based on random allocation of protocols of models A and B to the project communities. The unit of randomization will be the community. The list of communities will be provided by the project team and will include 56 locations in total of which 28 will be allocated to each of the models. The randomization will be stratified by district and allocated kit size. Identification of the program effects will be based on comparing the outcomes of interest between communities assigned to Model A and communities assigned to model B. The survey sampling frame will be based on 1) all farmers included in the irrigation and 2) a random sample of at least 10 farmers outside the kit yet inside the eligible area. The size of sample 2) will depend on how many of the randomly selected farmers are in fact included in the kit. In total around 1,400 farmers will be included – 25 in each community. By randomly sampling farmers in the eligible area we will establish the land distribution in the area and compare this to the distribution inside the kit. This will allow us to compare changes in probabilities of inclusion for smaller farmers between the two protocols. We will exploit the exogenous variation in irrigation exposure for smaller landholders in the eligible area to assess the impacts of irrigation. Additional details on computing treatment effects are described in Section 11.

It is possible that the placement of the infrastructure leads to redistribution and sales of land. Therefore we will careful track land ownership as soon as the location is determined. Household level data of the originally selected households will allow us to identify household level welfare impacts, whereas collecting plot level yield data of the current users of the irrigated plots will provide us with the yield impacts on the irrigated plots.

9.2 Sample Size Calculations

The randomization will be done at the cluster (the community) level. Our sample size is limited by the number of communities that will be treated and the number of farmers that will be surveyed in each of the communities. The project will build around 56 irrigation schemes. In each of the communities we will survey a random sample of in total of 25 farmers, resulting in a total sample of 1,400. Given these parameters on cluster size and the number of observations within each cluster, we calculate the minimum detectable effects (MDEs) assuming a power of 0.8 and an alpha of 0.1. Where there is no information available, we calculate them under potential scenarios of intra-cluster correlation (ICC). Below we present the power calculations for the main indicator directly targeted by the intervention:

proportion of small farmers in the eligible area that are selected. The final outcome indicator of interest is yield.

Proportion of small farmers

The main outcome indicator of interest is proportion of small farmers in the eligible area that are included in the scheme. We will select a random sample of 25 farmers within the eligible area to determine the distribution of cultivated land in the area. From the TIA data we see that in the region around 75% of farmers cultivate 2 or less hectares. We will compare the probabilities of these 19 small farmers to be selected under the two different selection protocols.

We don't have data on how correlated the selection of beneficiaries below the desired target land threshold will be within village. The proportion of households with less than 2 Ha of farm land who are selected to receive beneficiaries could have a high intra-cluster correlation at the village level if, for example, allocation of subsidies happens at the discretion of a single village chief in each village whose preferences vary mostly with regard to whether they prioritize smallest landholders or larger landholders. Alternatively, the intra-cluster correlation of the proportion of households below the cutoff who are selected may be low if villages use the preferences of many people to reach their allocations or if they use a diverse set of household characteristics to target. Without reliable information on ICC from these measures, we report the MDEs under a wide range of parameters shown below.

Departing from a scenario where under model B half of the small farmers are included, which yields the most conservative results, the MDE of an increase in small farmers being included ranges from 0.11 - 0.31.

ICC	MDES
0.05	0.11
0.10	0.13
0.15	0.15
0.20	0.16
0.30	0.19
0.40	0.21
0.50	0.23
1.00	0.31

Yield

Our sampling frame for yield will be based on the on average 15 households included in the scheme. To calculate the MDE on yield we use data from a Smallholders survey collected in an adjacent region of Mozambique (Kondylis, Mueller, and Zhu, 2015). Mean maize revenue per hectare (MZN/ha) is 1051, standard deviation 1186 and ICC 0.10. With a baseline and one follow up for yield outcomes we should be able to detect a 20.1% increase in revenue.

10. Data Collection

10.1 Quantitative Instruments

Proxy means data

The land size proxy is determined using a short survey, which includes 10 questions. The survey is applied during the selections visits by the project staff using tablets. The survey was elaborated and programmed by the IE team in collaboration with project management. The survey will be applied among all suggested beneficiaries in both model A and B communities. Only in model A villages will the program provide feedback on the result of the test. In addition we will collect the number of cultivated plots per household and self-reported total area cultivated.

Mapping of plots

One of the key objectives of the evaluation is to see if using strict targeting protocols using a proxy measure 1) can be used to successfully identity small farmers and 2) leads to a more equitable distribution of benefits, measured by the proportion of small farmers included. Therefore it is key to have an accurate measure of total cultivated land. The project team will take GPS coordinates of all plots of the farmers included in the kit, including those plots not covered by the kit, inside and outside the eligible areas. In addition, all plots of a random sample of at least 10 farmers within the eligible area will be mapped. Selection of plots will be done by randomly selecting grids in the eligible area and mapping all plots of the household cultivating on that location in the past 12 months or plots of the family cultivating the nearest plot. The mapping will be completed by the time the delivery of the kits is completed.

Maintenance and usage

Failure to properly maintain the kits or payment of fees for fuel will lead to slow deterioration or underutilization of the infrastructure. We are interested to see if different group compositions affects the ability for the necessary collective action. However, it is unlikely that the difference between the two models are so stark that we can observe differences by comparing binary outcomes such as if the scheme is working/being utilized after one year. Therefore we need measurements that allow us to identify differences on a more sensitive scale of possible deterioration. We will hire a data collection firm that will perform multiple visits throughout the year to measure efficiency of the pumps and the scheme by measuring water output for a given time at the head and tail end of the kit, respectively. In addition we will design a check list to identify compliance with most common maintenance activities based on the farmer training materials, which are currently being elaborated.

During the first year of operating, the project will cover the fuel costs of the pump. Only after the first year will we start collecting information on fee payments.

Agriculture surveys

Multi-module agriculture household surveys are planned for a sample of 25 farmers in all 56 communities that receive the irrigation kits. The surveys will capture relevant information to compute yield and profit such as self-reported landholdings, crop choice, harvest, sales and input use (labor, fertilizer, pesticides and self-reported water use) as well as general household characteristics and indicators of satisfaction with the selection procedure. The yield data will be collected separately for the irrigated and non-irrigated plots. Two large scale surveys are planned: the baseline survey is expected to go to the field in March 2016. This survey will allow us to identify the profile of selected versus general farmers in the area. The second survey is planned for March-2017. This data will allow us to identify the impact on yield of the irrigation in the different models. Funds will be raised to carry out a second follow-up after 2 years of operating.

10.2 MANAGEMENT OF DATA QUALITY

All data collection activities will be supervised by the Maputo based IE field coordinator in partnership with the SLWRMP/DPA team. The number of plots and land proxy data will be collected from different sources. The land proxy survey will be applied by extension agents that will be extensively trained by the IE and project team through classroom and field sessions. The model A selection procedure is more time consuming and local elite might be looking to game the system. The exact scoring algorithm is not shared with the agents to reduce this risk. In addition the agents will be aware that we will collect the same data in the baseline survey, which will be carried out by an independent firm and will allow us to cross-check the results. The mapping data will be triangulated with the number of self-reported plots from the proxy survey and baseline survey.

The agriculture data collection instruments will be piloted extensively in the field prior to going starting the data collection to ensure they are appropriate for the local context. Enumerators will participate in extensive training of the questionnaire and functioning of the tablets. Training will include classroom and field training. Enumerators will be selected based on their performance during the training. The data will be collected electronically, which allows us to program consistency checks and perform quality checks on a daily basis. Audits will be performed by recording parts of the interview and performing back-check interviews by a different team of interviewers. Cross-checking of the data will allow us to provide immediate feedback to the field teams in case of divergences or other problems.

10.3 Ethical Issues

All survey participants will be carefully informed about the data that will be collected throughout the study, the purpose of the surveys and the fact that their participation is voluntary. Only after participants provide consent will their data be collected. Strict protocols will be put in place to ensure data remains confidential. Any information that can link data to specific households will be removed after assignment of a unique identifier.

10.4 IE Implementation Monitoring System

The irrigation infrastructure is assigned to the communities and we expect excess demand. Therefore we are less affected by issues like delivery and take-up of the intervention. In contrast to large-scale irrigation interventions the kits require substantially less time to be placed – they kits are expected to be operational within 5 months of finalizing the beneficiary selection. Nevertheless the project team in collaboration with the field coordinator will carefully monitor all the different implementation phases from first community visits to delivery at each site, through a checklist with distinct outputs at each step. However, we are more concerned with the careful implementation of the selection procedures. Protocols to minimize these issues are described in section 10.4

11. Data Processing and Analysis

There are three primary types of outcomes we seek to explore:
1. Inclusion outcomes (at the individual and scheme level)
2. Contributions to the group operations and maintenance (at the individual and scheme level)
3. Yield and agricultural revenue improvements (at the individual level)

The inclusion outcomes at the individual level will be estimated by a binary dependent variable models which are a dummy variable for whether an individual household or plot was selected for inclusion in the scheme. Since the project’s stated goal is to include households who have no more than 2 Ha of cultivated land, we plan to divide the sample into the half of farmers who have less than or equal to 2 Ha of cultivated land and those who have more. We will then assess the relative strength of the two models for targeting by the following regression:

$$Y_{ij} = f\left(\alpha + \beta model\_A_j\right) + \varepsilon_{ij}$$

where Y_{ij} is a dummy for whether household *i* in village *j* was included and $model\_A_j$ is a dummy for assignment of village *j* to model A. f is the functional form of the binary dependent choice model used and could be logit, probit, or an LPM. This regression will be estimated separately by households who had more or less than 2 Ha of cultivated land at baseline to determine whether each targeting model is better at 1. Including more of the desired priority households and/or 2. Better at excluding more of the households meant to be excluded.

A second way of assessing whether plot size is differentially associated with inclusion across the two models is to estimate the following plot level regression:

$$Y_{ij} = f\left(\alpha + \beta_0 model\_A_j * area_{ij} + \beta_1 model\_A_j + \beta_2 area_{ij}\right) + \varepsilon_{ij}$$

where all $area_{ij}$ is the area of plot *i* in village *j*. The interaction of $model\_A_j$ and area tells us whether larger or smaller plots are differentially more likely to be included when the rule-based criteria is used.

The role of contributions to maintenance and operations the most difficult outcome to assess. The most straightforward way to measure differences in contributions across the model would be to estimate a regression of the form:

$$Y_j = \alpha + \beta model\_A_j + \varepsilon_j$$

where Y_j is a measure of total group contributions over a fixed time period such as contributions for fuel to run the irrigation pump or the total monetary value of group contributions to the maintenance fund over a fixed period. We plan to run these regressions and include Bonferroni multiple hypothesis testing corrections to account for the fact that we will be testing on a set of group outcomes for operations and maintenance. However, since we have a small number of villages, we do not expect to have a high degree of power for this set of tests.

One way of increasing power for the regressions concerning contributions to group operations and maintenance is to assess these outcomes at the individual level, since each individual will have to make contributions toward the fund. In the simplest case, such outcomes could be assessed by the simple regression:

$Y_{ij} = \alpha + \beta model\_A_j + \varepsilon_j$

where Y_{ij} is an outcome such as an individual's total contributions for fuel over a fixed time period. However, multiple issues arise with the interpretation of such a regression. The first is the issue that the number of group members is likely to vary by model. This means that there is no direct link between the size of each individual's contribution and the size of the total group contribution. The best way around this is to define the contributions in a standardized way for example, by considering an outcome such as contributions towards maintenance per Ha of area that household i cultivates under the scheme. Assessing this outcome in the above regression model assures that if farmers in model A villages are on average contributing more per Ha than farmers in model B villages, the total group fund must be receiving more contributions in Model A. The second issue with this model is that the characteristics of farmers included in the two models will differ by the model so that this regression will be able to identify differences in the contributions of farmers but not the mechanisms influencing contributions in the sense that Model B farmers may be shown to be giving higher contributions, but we will not be able to tell whether this happens because Model B leads to groups with smaller numbers of farmers or because Model B villages include beneficiaries who are wealthier.

Finally, we are ultimately interested in whether irrigation increases yields and whether the targeting method used to select users of irrigation mediates the magnitude of any potential gains. To assess the contribution of irrigation to yields, we can use two types of strategies. The first is a differences-in-differences (DID) model of the following form:

$Y_{ijt} = \alpha + \beta_0 irrigated_{ij} + \beta_1 post_t + \beta_2 post * irrigated_{ijt} + \varepsilon_{ijt}$

where Y_{ijt} is a measure of agricultural output such as yield per Ha or revenue per Ha, $irrigated_{ijt}$ is a dummy for whether the household was selected to receive the kit, $post_t$ is a dummy for whether the observation comes from the season following the installation of the kit, and $post * irrigated_{ijt}$ is the interaction of post and irrigated. The ATE of the irrigation on yields is identified by β_2. The identifying assumption here is that the only differences between the households selected or not to receive irrigation are level differences accounted for β_0 irrigated. Endogenous selection of beneficiaries of the irrigation equipment will bias the estimated effect of irrigation if communities anticipate that some farmers will be able to improve their yields even in absence of the irrigation and allocate the equipment to these households.

In order to limit endogeneity concern about described above, we can also exploit the rules in the rule-based selection to estimate the LATE of irrigation on small landholders in two ways. First, we use the fact that small land holding should predict greater inclusion rates in the irrigation only for households living in villages randomly selected into model A. This fact allows us to estimate the following IV regression:

$$\Delta Y_{ij} = \alpha + \beta_0 irrigated_{ij} + \beta_1 small_{ij} + \beta_2 model\_A_j + \varepsilon_{ijt}$$

where ΔY_{ij} is the change in agricultural revenue from t = 0 to t= 1 for household i, $small_{ij}$ is a dummy for whether the household cultivated less than or equal to 2 Ha at baseline, and $model\_A_j$ is a dummy for whether the village was randomly assigned to model A. $irrigated_{ij}$ is a dummy that equals 1 if the household receives the irrigation kit and is instrumented by the interaction of $small$ and $model\_A$.

The second way to exploit the rules in Model A to is to employ an RD design. The proxy means test uses a sharp cutoff to determine whether a household is eligible or not. Although households below the cutoff and households above the cutoff are likely to be very different, households just above and just below the cutoff are likely to be similar. Since the scores are computed with a tablet that reports only the result of the test and not the components, households will only see the result of their eligibility test and will not know whether they were close or far from the cutoff or be able to easily manipulate the score to get just above the score. To implement the RD design we will estimate the following regression:

$$\Delta Y_{ij} = \alpha + \beta_0 irrigated_{ij} + \beta_1 score_{ij} + \varepsilon_{ijt}$$

Where all variables are defined as before, with the addition of $score_{ij}$ which is the score assigned to the household by the proxy means test. This regression will be estimated only among Model A villages for whom the rules apply and will be reported by limiting the sample to gradually increasing bandwidths of scores around the eligibility cutoff. Bigger bandwidths will increase the number of observations, but increase the endogeneity concerns related to comparing households whose scores are increasingly different.

For all individual regressions, standard errors will be clustered at the village level, which is the unit of randomization. In addition, fixed effects for the size of the kit assigned to the village (5 Ha or 10 Ha) will be included to account for the level of stratification in the randomized assignment.

We plan to register the evaluation and pre-analysis plan at either the AEA RCT Registry or 3ie Registry.

12. Study Limitations and Risks

Like with most impact evaluations and field experiments, there are a number of limitations and risks. Where possible we will minimize such risks.

The primary threat to internal validity arises from the relatively small number of units to be randomized into treatment and control. With only 56 villages included, there is a risk that randomization will not achieve balance on observable and unobservable characteristics of farmers that may be associated with their outcomes. We plan to address this risk by stratifying the randomization on district and size of the scheme. We also plan to conduct balance checks on a set of pre-specified variables to assess the degree of balance and will also present regressions that control for variables that show important imbalance in these balance checks. Second, communities could fundamentally change because of the cooperative process involved in deciding where the kit will be placed (in community targeting group) and that could result in confounding of the core comparison – communities may become more cooperative as a result of the site selection process and so we would capture the effect of both community targeting AND community cohesion. We will limit this concern in a few ways. First, in our baseline, we will collect community level characteristics and see the evolution of these between the two groups over the course of our study. Second, we recognize that if community targeting of a public good with positive spillovers is resulting in community cohesion, then that is an added benefit that should be documented in our impact evaluation.

The primary threats to external validity arise from whether the irrigation schemes and farmers groups selected are representative of the type of populations who would be involved in similar interventions if they were to be employed elsewhere. However, given that this is an active program being run by the national government rather than only a pilot project designed for research purposes alone, it is likely that similar selection processes occur whenever governments involved. Therefore, it is likely that the groups and locations ultimately selected for this intervention are similar to those who would end up participating if the project were planned again.

13. Plan for Using Data and Evidence from the Study

We will be actively involved in the dissemination of evidence acquired during the course of this process to policy makers, practitioners and academics. First, we have been actively involved with policy makers at different levels during the design phase. Baseline and monitoring data will help the provincial administration during the program implementation of the project as well as inform the broader irrigation investments made in the country at the national level. Upon completion of the evaluation we will work closely together with all stakeholders to elaborate relevant policy briefs and dissemination events. A report will be produced by the research team to be shared with project staff, the TTL's and policy makers from relevant departments to summarize learning, solicit suggestions and improvements, and generate new uses for the resulting data.

Second, through the global DECIE network we are working closely with different stakeholders in the development arena. The network brings together governments, TTLs from different MDBs, multiple donors and academics. The IE was proposed and designed with participation from TTL's, project staff and the core research team and a subset of that group met in Kigali in June 2014 as part of DECIE's broader initiatives in the areas of agriculture and food security. The results will be disseminated widely across the community of practice through the annual workshops as well across the irrigation projects specifically through close collaboration between the research teams.

In addition, we plan to make our finding broadly available to other WB and independent agriculture and irrigation related projects to emphasize the role of community targeting of public goods. Finally, we plan to develop a series of ambitious research papers from the experiment and the results and engage the broader academic community to both contribute to and shape the knowledge from this IE. We hope that such academic work is widely regarded in seminars and conferences and eventually published in an academic economics or general interest journal of the top caliber.

All data will be made available online through the IE database, following the Bank's open data policy.

REFERENCES

Acemoglu, D., Reed, T. and Robinson, J.A. (2013): 'Chiefs: Economic Development and Elite Control of Civil Society in Sierra Leone', *NBER Working Paper Series*, Working Paper 18691

Arndt, C., Hussain, A., Jones, S., Nhate, V., Tarp, F. and Thurlow, J. (2012): 'Explaining the Evolu1on of Poverty: The Case of Mozambique,' *American Journal of Agricultural Economics* 94 (4), pp. 854-872

Alatas, V., Banerjee, Hanna, R., Olken, B., Purnamasari, R. and Wai-Poi, M. (2013): "Does Elite Capture Matter? Local Elites and Targeted Welfare Programs in Indonesia', *NBER Working Paper Series*, Working Paper 18798

Alatas, V., Banerjee, A., Hanna, R., Olken, B and Tobias, J. (2012): 'Targeting the Poor: Evidence from a Field Experiment in Indonesia', *American Economic Review* 102(4), pp. 1206–1240

Besley, T. and Coate, S. (1992): 'Workfare versus Welfare: Incentive Arguments for Work Requirements in Poverty-Alleviation Programs', *American Economic Review* 82(1), pp. 249–61

Basurto, P., Dupas and Robinson, J. (2015): 'Decentralization and Efficiency of Subsidy Targeting: Evidence from Chiefs in Rural Malawi', Working Paper

Cunguara, B. and Kelly, B. (2009): 'The Impact of the PARPA II in Promoting the Agricultural Sector in Rural Mozambique,' A Study as Input to Impact Evaluation Report (RAI) of PARPA II, Maputo, Mozambique

Duflo, E. and Pande, R. (2007): 'Dams', *The Quarterly Journal of Economics*, pp. 601-646

Dupas, P., Hoffman, V., Kremer, M. and Zwane, A.P. (2013) 'Micro-ordeals, targeting, and habit formation

FAO (2015): FAOSTAT Online Country Profiles - Mozambique, Italy, Rome
[Accessed via: http://faostat.fao.org/CountryProfiles/Country_Profile/Direct.aspx?lang=en&area=144]

FAO (2005): Irrigation in Africa in figures - AQUASTAT Survey 2005

Hoffman, V., Jakiela, P., Kremer, M. and Sheely, R. (2015) : 'Targeting, Discretionary Funding, and the Provision of

Local Public Goods: Evidence from Kenya’

Isaac, R.M, Walker, J.M and Williams, A.W. (1994): ‘Group size and the voluntary provision of public goods – experimental evidence utilizing large groups’, *Journal of Public Economics* 54, pp. 1-36

Jack, K. (2013): ‘Private Information and the Allocation of Land Use Subsidies in Malawi’, *American Economic Journal: Applied Economics* 5(3), pp. 113–135

Jones, S. and Tarp, F. (2012): ‘Jobs and Welfare in Mozambique: Country Case Study for the 2013 World Development Report,’ UNU-WIDER, Helsinki, Finland

Karlan, D. and Thuysbaert, B. (2013): ‘Targetting Ultra-Poor Households in Honduras and Peru’, *NBER Working Paper Series*, Working Paper 19646

Keefer , P. and Khemani, S. (2009): ‘When Do Legislators Pass on Pork? The Role of Political Parties in Determining Legislator Effort’, American Political Science Review 103(1), pp. 99-112

Kondylis, F., Mueller, V. and Zhu, J. (2014) : ‘Seeing is Believing? Evidence from an Extension Network Experiment’, *Policy Research Working Paper Series*, Working Paper 7000

Lizzeri, A. and Persico, N. (2001): ‘The Provision of Public Goods Under Alternative Electoral Incentives’, The American Economic Review 91(1), pp. 225-246

Mather, D., Cunguara, B. and Boughton, D. (2008): ‘Household Income and Assets in Rural Mozambique, 2002-2005: Can Pro-Poor Growth be Sustained?’, Research Paper Series, Research Report 66, Mozambique Ministry of Agriculture, Maputo, Mozambique

MINAG (2013) 'Estratégia de Irrigação,' Ministry of Agriculture, Maputo, Mozambique

MINAG (2012) 'Inquérito Agrícola Integrado, Ministry of Agriculture, Maputo, Mozambique.

MINAG (2010) ‘Plano Estratégico para o Desenvolvimento do Sector Agrário (PEDSA),’ Ministry of Agriculture, Maputo, Mozambique

Smolensky, E, Tideman, T. N. and Nichols, D. (1971): ‘ ‘, *Papers in Regional Science* 26(1), pp. 37–52

Nichols, A.L. and Zeckhauser, R.J. (1982): ‘Targeting Transfers through Restrictions on Recipients, *The American Economic Review* 72 (2), pp. 372-377

Olson, M. (1965): ‘ *The Logic of Collective Action: Public Goods and the Theory of Groups’,* Harvard University Press: Cambridge, MA, 1965.

Ostrom, E. (2003): ‘How Types of Goods and Property Rights Jointly Affect Collective Action’, *Journal of Theoretical Politics* 15(3), pp. 239-270

Ravallion, M. (1991): ‘Reaching the Rural Poor Through Public Employment: Arguments, Evidence, and Lessons from South Asia’, *World Bank Research Observer* 6(2), pp. 153–175

ANNEX A – CONSTRUCTION OF SCORECARD

The project agreed that they wanted to prioritize small landholders, which they agreed after discussion with extension staff and comparison with the TIA data would be those with less than or equal to 2 Ha of land that they actively cultivate (as defined by cultivation in the last 12 months). The methods that follow are heavily based on Shreiner and Lory (2013). That paper describes the construction of a poverty scorecard for Mozambique by a specialist on the subject.

To identify proxies associated with owning less than or equal to 2 Ha, we use data from the national representative agriculture survey TIA. We restricted the data to households that live in Gaza province and those whose primary activity is farming to most closely match our sample.

Consistent with the machine learning literature, we first randomly select half of the data. The first half is assigned to be "Training Data" and the second half is "Testing data". The idea is to use the testing data to fit a model and then use the training data to show and assess how well it fits. Using all the data may yield a slightly tighter fit, but prohibits the user from having any kind of reasonable assessment of fit. This left us with approximately 230 households in each set.

To generate the model, we selected household characteristics available in the data. These were chosen both by comparing the list used in Shreiner and Lory and choosing additional items that could be relevant. Once the candidate variables are selected, we ran the following algorithm:

1. Run a logit model for predicting a binary variable for a household meeting the land size target on each of the candidate variables one at a time.
2. From each logit regression, measure the goodness of model fit from the area under the LROC curve.[1]
3. Choose the model that has the highest LROC. Designate the variable for that model as the first selected variable.
4. Re-run logits for the target on the selected variable and each of the remaining candidate variables. Designate the model with the highest measure of goodness of fit as the second candidate variable.

[1] The LROC summarizes the percentage of cases in which various rules based on the logit model would correctly identify the household as successfully meeting the target.

5. Repeat step 4 adding each candidate variable one at a time and trying each of the remaining candidate variables until 14 variables have been selected.

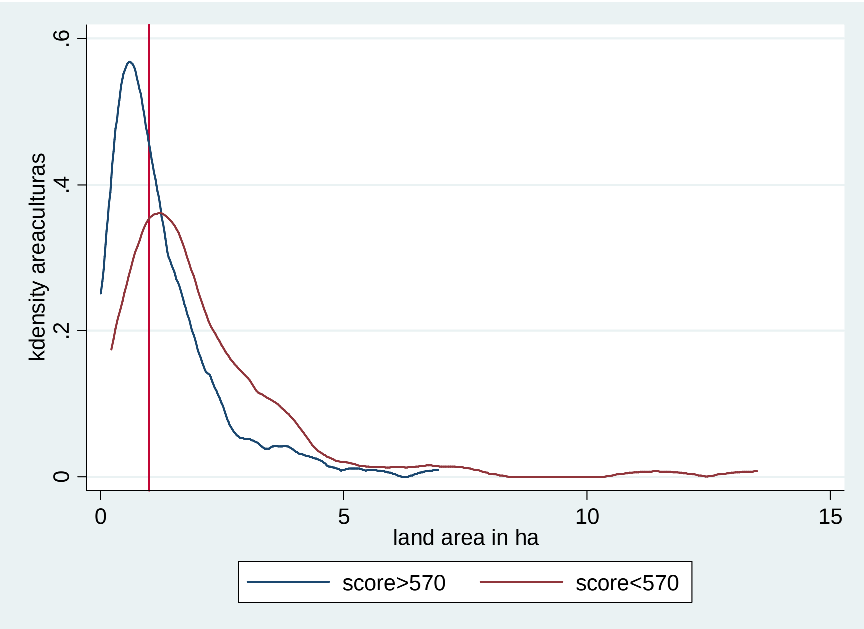
The figure below shows the LROC curve for the model selected on the training data (the data used to fit the model). It doesn't look to bad, suggesting that these variables do a decent job sorting out the households.

For the construction of the score card we converted the logit coefficients to more comprehensible score with the following features:

- They should all be positive
- They should all be integers with at most 2 digits
- For any given variable, the weight should be intuitive, straightforward in its association to landsize, and not confusingly contradictory in its association with landsize within the same variable.

To assess the scorecard we look how well it does in successfully predicting outcomes in the training data. Recall that the training data is the 50% of households kept out when fitting the model. If we run the scores on these households, we get an idea for how the model performs on real households that come from the same location but were not used to generate the model.

The following graphs show one way to assess how the model performs. The graph below shows the land distribution for households above and below the median score. Clearly households above the median score have smaller land sizes and those with scores above the median have larger scores, which is what we want. Most of the failed cases were close to the cutoff, so it's not the case that the biggest landholders are getting through and the smallest are getting rejected. The cut-off was set at 570, the median score. In the testing data with 570 as the cutoff, both the inclusion and exclusion error are just under 30%.



ANNEX B – PRIORITY TEST SCORE CARD

This version of the score card is currently being piloted and will likely suffer alterations.

Question	answer	Score
Do any members of your household do any own, not salaried, work to produce a good besides agriculture?	Yes	59
	No	0
Do any members of your household do salary work?	Yes	52
	No	0
How many members live in your household?	1 to 2	114
	3 to 4	121
	5 to 7	49
	More than 7	0
Does your household own any chickens?	Yes	0
	No	110
Do you own a granary?	Yes	0
	No	115
Do you practice line planting?	Yes	0
	No	96
Does your household own a mobile phone?	Yes	0
	No	137
Does your household own an axe that you used for farming in the last 1	Yes	0

year?	No	75
Does your household own any sheep?	Yes No	0 84
How many members of your household own their primary income from agriculture?	1 2 More than 2	66 23 0
How many cattle does your household own?	0 1 to 10 11 to 20 more than 20	100 105 29 0
How many goats does your household own?	0 1 2 more than 2	38 86 19 0

14. IE Management

14.1 Evaluation Team and Main Counterparts

TABLE 2 – IE TEAM AND MAIN COUNTERPARTS

Name	Role	Organization/Unit
Florence Kondylis	Principal investigator, IE TTL	WBG/DECIE
Paul Christian	Principal investigator	Cornell University
Teevrat Garg	Principal investigator	LSE, UCSD
Astrid Zwager	Principal investigator, IE Coordinator	WBG/DECIE
Steven Glover	IE Field Coordinator	WBG/DECIE
Olagoke Oladapo	Chief Agro Economist, Project TTL	Afdb
Francisco Matuca	M&E officer	DPA
Paiva Munguambe	Director	INIR

14.2 WORK PLAN AND DELIVERABLES

TABLE 3 – MILESTONES AND TIMELINE

Milestones	Completion Date
Peer-reviewed Concept Note	September 2015
Identification of locations and beneficiaries	September – November 2015
Mapping of plots	September 2015 – March 2016
HH Data collection (Baseline)	March 2016
Delivery if irrigation infrastructure	March 2016
First data analysis	June 2016
Monitoring visits	March 2016 – March 2017
HH Data collection (Follow-up)	March 2017
Final report and policy notes	August 2017
Dissemination of findings	September 2017