Impact Evaluation Designs for the Mozambique-MCA Land Project: Improving Site-Specific Access to Land Activity in Urban and Rural Areas (Activities III), and the Institutional Strengthening of the Land Administration System (Activities II)

Submitted to the
Millennium Challenge Corporation

By

Michigan State University (MSU) and the Ministry of Agriculture Department of Economics (MINAG-DE)

Songqing Jin, Mywish Maredia, Raul Pitoro, Gerhardus Schultink, and Ellen Payongayong

October 11, 2016

ABSTRACT: The Land Tenure Services Project (or simply the *"Land Project"*) of the Mozambique MCA compact aims to establish more efficient and secure access to land by improving the policy and regulatory framework and helping beneficiaries meet their immediate needs for registered land rights and better access to land for investment. The Land Project consists of three main types of activities (Activities I, II and III) and several component activities that will be implemented at different levels of geopolitical aggregation (i.e., national, provincial, District, Municipal, "hot spots" areas, etc.). The Michigan State University (MSU) was contracted by MCC to evaluate the impacts of different activities under the Land Project. Specifically, MSU is responsible for evaluating the activities that can be rigorously evaluated, namely, the activities that allow us to establish valid treatment and control groups. This document lays out three separate evaluation designs that evaluate different activities under the Land Project. They include (1) evaluation of the institutional strengthening activity (Activity II) which involves upgrading land administration system in selected municipalities and districts, (2) evaluation of the 'hot spots' (or site-specific) activities (Activity III) in urban areas and (3) evaluation of the 'hot spots' activity (Activity III) in the rural areas. For each of the evaluation designs, we proposed evaluation method(s), survey and sampling strategies, and times for the baseline and the endline surveys. At the time we worked on this version of the design document, we have already collected the baseline data for all the three evaluations. For the 'hot spot' activities in both the rural and urban areas, the fact that the actual intervention did not follow the original implementation plans could affect the validity of the original evaluation designs. Thanks to the baseline survey data, we were able to recalculate the statistical powers and use them to reassess the validity of the original designs. Recommendation on whether evaluation for a particular activities should be pursued or not are made based on the revised statistical power.

Contents

1. Overview of the Land Project 1

2. Impact Evaluation of the Land Project: An Overview 2

2.1. Overview of the activities being evaluated 2

2.2. Overview of the project areas 4

2.3. Evaluation approach 5

3. Impact Evaluation Design of the *Institutional Strengthening Activities* (ISA) of the Land Project 7

3.1. Overview of the activity being evaluated and expected impacts 7

Table 3: Impact pathway for the Institutional Strengthening Activities (ISA) of the Land Project 7

3.2. Geographic coverage of the IE and identifying comparison group 8

3.3. Identification issues 9

3.4. Data collection 10

3.5. Some useful lessons from the baseline data 11

4. Impact Evaluation Design of the Urban Hotspot Activities Under Activity III of the Land Project 12

4.1. Overview of the activity being evaluated 12

4.2. Geographic coverage of the IE 12

4.3. Research questions to be addressed by the IE 13

4.4. Identifying Comparison Communities 14

4.5. Sample size and Power Calculation 15

4.6. Sampling Method 17

4.7. Data collection 19

4.8. Assessing the original design after project implementation (as of May 2016) 19

4.8.1. Validity of the original design for Monapo vila 19

4.8.2. Validity of the original design for the Nampula City 20

4.8.3. Justification of MDEs for non-randomization and additional co-variates 22

4.9. Implications for next steps on evaluation of Activity III in Nampula 22

5.1. Overview of the activity being evaluated 23

5.2. Geographic coverage for the impact evaluation 23

5.3. Research questions addressed by the IE 24

5.4. Identifying the comparison communities 24

5.5. Sample size and sample selection 25
5.6. Power Calculation 29
5.7. Baseline data collection 31
5.8. Assessing the Validity of the Original Evaluation Design (As of May 2016) 33
5.8.1. Validity of the Evaluation Design for the Malema District 33
5.8.2. Validity of the Evaluation Design for the Mecufi District 33
5.9. Implications and Next Steps on Evaluation of Activity III in Malema and Mecufi 34
References 36
Annex 1: Definition of important terms and concepts in the context of the Land Project 37
Appendix Table 1: Intra-cluster correlation coefficients (ρ) of urban communities based on baseline data collected in 2011 38
Appendix Table 2: Intra-cluster correlation coefficients (ρ) of rural communities based on baseline data collected in 2011/2012 39

Impact Evaluation Design for the Mozambique-MCA Land Project: Improving Site-Specific Access to Land Activity in Urban and Rural Areas, and the Institutional Strengthening of Land Administration System

1. Overview of the Land Project

The Government of the Republic of Mozambique and the Millennium Challenge Corporation (MCC), on behalf of the United States Government, signed a Compact Agreement (which entered into force on September 22, 2008) for a US $507 million grant to be implemented over a 5-year period. The overall objective of the proposed Program is to reduce poverty through economic growth in the four Northern Provinces of Mozambique (Niassa, Cabo Delgado, Nampula, and Zambézia). The Program involves crucially needed investments in water, sanitation, and transport infrastructure, land tenure security, agriculture, capacity building, and institutional strengthening.

The Land Tenure Services Project (or simply the *"Land Project"*) of the Mozambique MCA compact aims to establish more efficient and secure access to land by improving the policy and regulatory framework and helping beneficiaries meet their immediate needs for registered land rights and better access to land for investment. The Project's objectives are to: (i) increase the level and value of investment on land; (ii) increase access to land; (iii) reduce the costs associated with acquiring land user rights; and (iv) resolve and prevent conflicts over land. Investments are targeted to all four Northern Provinces, at all levels of administration – National, Provincial, and District / Municipal – and across a range of beneficiaries, including rural individual land holders, rural communities, urban land holders, and domestic and international investors.

The Land Project consists of three main types of activities and several component activities that will be implemented at different levels of geopolitical aggregation (i.e., national, provincial, District, Municipal, priority/"hot spots" areas, etc.). Overall, the *Land Project* works on improving policy, upgrading public land administration agencies (the title registry and cadastre), and facilitating site-specific land access. The three main types of activities described above (Activities I, II and III) address concerns widely shared across the private sector, government, and civil society with solutions that bring together their diverse perspectives. Benefits from the Land Tenure Services Project are projected to accrue to (i) rural households; (ii) urban households; (iii) communities; and (iv) business enterprises and investors in the form of increased income, lower transaction costs, and greater investment opportunities.
The specific activities are described below:

Policy Monitoring Activity (Activity I) (all activities under Activity I implemented at the **national level**): Support for an improved policy environment, including addressing implementation problems for the existing land law and engaging in regulatory review to improve upon it. Examples of activities include:

1. Development of a national land administration regulatory framework and needs assessment
2. Formation of Land Policy Consultative Forum that will provide technical and logistical support to monitor progress on land legislation reform and implementation. It is worth noting

that little occurred beyond needs assessment and holding of the land policy consultative forum.[1]

3. A broad campaign of public education, outreach and increasing awareness of non-judicial dispute resolution methods
4. Expand program on legal and judicial training to paralegals
5. Advisory services to DNTF

Capacity Building Activity (Activity II): Building the institutional capacity to implement policies and provide quality public land-related services. Examples of activities under Activity II include:

1. Development of LIMS (**national level**)[2]
2. Professional development and training, and upgrading of facilities (**4 Provincial** SPGCs, 8 municipal and **12 selected district** land service offices)
3. Technical Assistance for cadastral development in selected municipalities (**8 selected municipalities).**

Site-specific Activity (Activity III): Facilitating access to land use by helping people and businesses with (i) clear information on land rights and access; (ii) resolution of conflicts with more predictable and speedy resolution of land and commercial disputes – which in turn creates better conditions for investment and business development; and (iii) registering their grants of land use (land titles to long-term or perpetual-use rights). Examples of activities include:

1. Mapping and right inventory exercise (all **12 selected districts and 8 municipalities**) and piloting an approach to area-wide registration of land rights in "Priority areas"; Streamlining investor and farmer access to land by making available simple informational tools and guidelines (**selected hotspot areas** within the 12 districts and 8 municipalities)
2. Support of the Community Land Fund (iTC) (**3 provinces** –Zambezia, Nampula and Niassa). Initially established by a coalition of donors and implemented in Inhambane, Cabo Delgado, and Manica provinces, in 2009 it was replicated and funded by the Land component of MCA to support the community land delimitation, registration, negotiations, and resource planning (MINAG, 2011c).

2. Impact Evaluation of the Land Project: An Overview

2.1. Overview of the activities being evaluated

As described above, the Land Project consists of three main types of activities and several component activities that will be implemented at different levels of geo-political aggregation (i.e., national, provincial, district, municipal, priority/"hot spots" areas, etc.). Because of different

[1] Legislation such as ability to transfer a DUAT was never changed, which led to continued difficulties in transferring land rights.

[2] **The LIMS system was initially planned for installation in DNTF, 4 SPGCs and 8 municipalities. It was later extended on the basis of a needs assessment conducted in January 2013 to another 6 SPGCs. So ultimately LIMS was installed with MCC funds in all 10 provinces in the country. The role of other donors was to support later stages of LIMS operationalization – i.e. training and other activities for which there was insufficient time to be completed with MCC funds due to the late date of LIMS finalization and installation, particularly in the 6 additional provinces.**

geographic scale and diverse scope of activities across selected provinces, districts and municipalities, it is not possible to implement a rigorous impact evaluation of the Land Project as a whole. Thus, a multi-faceted evaluation approach is designed to assess the short- to medium term impacts of the Land Project across the three 'Activities' (Activities I, II and III):

Activities under Activity I and Activity II: The coverage and scope of project activities under Activities I and II (capacity building and policy monitoring) range from national to provincial to district/municipal level. An impact evaluation of the project activities under Activities I and II is meaningful only if the evaluation can identify the causal effect of the activities on project outcomes. To establish a causal relationship between project activities and project outcomes requires both the treatment group (areas that are affected by the activities) and control group (areas that have not been affected or will not be affected by the project activities in the duration of the project evaluation exercise). For this reason, all activities under Activity I cannot be rigorously evaluated because there will be no control group for these national level activities. However, for the activities that will be implemented at the provincial level or the district/municipal level to strengthen their respective land administration systems (or simply "institutional strengthening activities"), potential evaluation is feasible with a proper design, and sufficient and appropriate data. Specifically, the causal effects can be identified by comparing outcomes in provinces (municipalities/districts) where the institutional strengthening activities were implemented (the "treatment group") to the outcomes in other provinces (municipalitiest/districts) where those activities have not been or will not be implemented (the "control group"). Essentially all activities under Activity II fall into this category.

Activities under Activity III: Impact evaluation (IE) of the land project activities implemented under Activity III will focus on "hotspot" issues in selected priority areas that result in registering or granting land use rights (i.e., land titles to long-term or perpetual-use rights) to individual households.[3] The hotspot activities will be implemented in selected urban areas and rural areas. Two separate evaluations were designed to evaluate the urban and the rural hotspot activities, separately. The two evaluations of the activities under Activity III were designed as semi-rigorous impact evaluations. Unfortunately, for various practical reasons, the implementation of Activity III ended up treating some areas that were intended as controls and failed to reach some areas that were intended as reatment areas.[4] As a result, it was necessary for us (the MSU evaluation team) to reassess the validity of the original designs and propose alternative designs.

This design document is comprised of three designs to evaluate three separate activities of the Land Project – an impact evaluation of site-specific activities in urban hotspot areas, an impact evaluation of site-specific activities in rural hotspot areas and an impact evaluation of institutional strengthening activities. It is important to note that the areas receiving site-specific activities in urban or rural hotspot areas are also affected by all the activities under Activity I and Activity II. Both urban and the rural site-specific evaluations would essentially evaluate the effects of receiving Activities I, II and III versus just receiving I and II in urban and rural areas. Similarly,

[3] Initially, there were plans to conduct rigorous IE of the community land fund project (iTC) under Activity III. However, based on the design of the iTC project and given the vast and diverse issues to be potentially covered by iTC, it was not feasible to do a rigorous impact evaluation of this component of Activity III.

[4] In fact, all the control areas in Monapo municipality were treated, so no valid impact evaluation for Monapo municipality is feasible. On the other hand, the original impact evaluation design also undermined the one for Nampula municipality, as some treatment areas in the Nampula municipality were not treated or only partially-treated. Whether a rigorous impact evaluation of Activity III in Nampula municipality is possible requires recalculation of the statistical power.

the evaluation of institutional strengthening activities would essentially evaluate the effects of receiving institutional strengthening activities plus all the other activities under Activities I and II versus only receiving all the other activities under Activities I and II without the provincial, municipal or district level institutional strengthening activities.

2.2. Overview of the project areas

The Land Project targeted 8 municipalities and 12 districts in four Northern Mozambique provinces. A list of all the 8 municipalities (Quelimane, Mocuba, Monapo vila, Nampula city, Pemba, Mocimboa da Praia Vila, Lichinga, and Cuamba) and the selection criteria they meet for Land Project activities is given in Table 1. A list of all the 12 districts (Nicoadala, Morrumbala, Mocuba, Malema, Monapo, Moma, Mocimboa da Praia, Montepuez, Mecufi, Majune, Lichiga, and Metangula) and the selection criteria they meet for Land Project activities is given in Table 2. The four Northern provinces are Zambezia, Nampula, Cabo Delgado and Niassa. While the institutional strengthening activities are expected to affect the land administration systems in all the 8 municipalities and 12 districts, the urban and rural site-specific activities will be implemented in some prioritized areas in the 8 municipalities and 12 districts.

Table 1: Selection criteria met by the eight municipalities selected for Land Project activities in four Northern provinces

	Criterion 1	Criterion 2	Criterion 3	Criterion 4	Criterion 5	Criterion 6	Group
Zambézia Province							
Quelimane-cidade			X			X	G
Mocuba-cidade		x	X		X	X	F
Nampula Province							
Monapo-Vila	x	x	X		X	X	E
Nampula-cidade	x		X		X	x	D
Cabo Delgado Province							
Pemba-cidade	x	x			X	x	C
Mocimboa da Praia- Vila	x	x				x	B
Niassa Province							
Lichinga-cidade	x	x	X			x	A
Cuamba-cidade	x	x	X			x	

Key for Criteria: 1 = high demand for DUATs; 2 = government priority; 3 = local technical capacity exists; 4 = support from other sources (financial and human); 5 = land use plans exist; 6 = high risk of land conflicts.

Table 2: Selection criteria met by the 12 districts selected for Land Project activities in four Northern provinces

	Criterion 1	Criterion 2	Criterion 3	Criterion 4	Criterion 5	Criterion 6	Group
Zambézia Province							
Nicoadala	X	X				X	I
Morrumbala	X	X	X			X	II
Mocuba	X	X	X			X	
Nampula Province							
Malema	X	X	X			X	III
Monapo	X	X	X			X	

Moma	X	X				X	VI
Cabo Delgado Province							
Mocimboa da Praia	X	X		X	X	X	IV
Montepuez	X	X		X	X	X	
Mecufi	X	X		X		X	V
Niassa Province							
Majune	X	X				X	I
Lichinga	X	X				X	
Metangula						X	VII

Key for Criteria: 1 = high demand for DUATs; 2 = government priority; 3 = local technical capacity exists; 4 = support from other sources (financial and human); 5 = land use plans exist; 6 = high risk of land conflicts.

2.3. Evaluation approach

By conducting an impact evaluation of the different activities under the Land Project (i.e., site-specific hotspot activities in urban area, site-specific activities in rural area, and institutional strengthening activities), we intend to quantitatively estimate the change in population attributes that is attributable to the implementation of the relevant activities under the Land Project. Thus we plan to compare the outcomes of the targeted population in the presence of the program relative to the population's outcomes if the program had not been implemented. In other words, the basic principle that guides our approach is the comparison between situations "with" the project activities and "without" the project activities. This is as opposed to merely comparing beneficiaries "before" and "after" the project implementation. Unfortunately, it is not possible to compare the same population simultaneously under both conditions --with and without program exposure, because a given household or community (depending on the unit of intervention) is either treated or not, but not both.

Practically, to address this problem, we estimate the average impact of the program on a group of individuals by comparing them to a similar group of individuals that are not directly affected by the program. Therefore, one critical step of any impact evaluation exercise is to establish a credible control group. A number of different empirical approaches have been employed to establish the credible comparison group (or control group). The most robust approach is randomization – in which the treatment group and control group are randomly selected from all eligible sampling units (either clusters or individuals). A randomized experiment guarantees that (on average) there are no differences in the observed and unobserved characteristics between the treatment and control group and thus, a statistically significant difference in outcomes between the two groups is attributed to the program.

While the "gold standard" of impact evaluation is randomization control trial (RCT), this is not always possible in practice. For example, the 8 municipalities and 12 districts to receive institutional strengthening activities as well as the prioritized urban areas in the 8 municipalities and prioritized rural areas in the 12 districts to receive site-specific activities, are not randomly chosen. In fact, these areas were pre-selected by national or local governments to receive these activities for economic development or other practical reasons. Given the non-random selection of program areas for all the three types of activities, we have to use (an) alternative evaluation approach(es) to evaluate the institutional strengthening activities and the urban and rural site-specific hotspots activities. Specifically, we will use the difference-in-difference (DID) approach for all the three separate evaluations.

The DID approach essentially measures the difference of outcome indicators between participants (treatment group) and nonparticipants (comparison group) before and after program intervention. In the context of panel data (with a baseline survey and a follow up survey of the same communities or households), DID is a common and valid method to estimate the impact of an intervention if the assumption holds that unobserved heterogeneity is time-invariant and uncorrelated with the treatment effect. While the main advantage of DID is its ability to control for time invariant unobserved factors, its assumption of constant selection bias over time may be unrealistic in practice. [5]

Let Y be the outcome of interest (e.g., total number of DUATs issued or the average time lapse between application and issuance of a DUAT in the case of institutional strengthening intervention, or land investment, land market participation, household income, off-farm employment in the case of site-specific intervention of DUAT issuance). Our goal is to evaluate the impact of a specific intervention $\boldsymbol{T}$ (i.e., upgrading of the land administration system in the case of institutional strengthening activities, or issuing DUATs to urban or rural residents in the case of site-specific activities) on Y after a time period 1. Specifically, we can achieve this evaluation through DID as:

$$DID = E[Y_1^T - Y_0^T] - E[Y_1^C - Y_0^C] \qquad (1),$$

where the superscripts T and C refer to treatment and control units (municipality or district in the institutional strengthening activities, or households in the two site-specific interventions), respectively; the subscripts 1 and 0 refer to time period 1 (after the intervention) and time period 0 (the baseline period), respective; T=1 refers to Treatment group. The regression counterpart of (1) is the following:

$$Y_i = \alpha + \beta T_i + \gamma t + \delta(T_i * t) + \varepsilon_i \qquad (2),$$

Where T_i is the dummy to distinguish treatment group (T=1) from control groups (T=0), t is a time dummy (t=0 for before treatment and t=1 for after the treatment). In (2), we can further add other control variables ($\boldsymbol{X}$) to increase the efficiency of the estimation. DID is widely used in impact evaluation of policy interventions especially when the RCT-based data are not available (see discussion by Duflo, Glennerster and Kremer 2007; Ravallion 2005). The DID approach was also used by similar studies on land titling projects in other countries (Deininger et al. 2011, Di Tella 2007; Field 2007).

[5] For the rural site-specific and urban site-specific hotspot activities, we can also combine the DID with the propensity score matching (PSM) method to further improve the reliability of the estimated impacts.

3. Impact Evaluation Design of the *Institutional Strengthening Activities* (ISA) of the Land Project

3.1. Overview of the activity being evaluated and expected impacts

The institutional strengthening activities (ISA) to be evaluated are a subset of activities under Activity II that include professional development and training, upgrading facilities, and assistance to the development of municipal and district LIMS in selected municipalities and districts in 4 Northern provinces (column 1 of table 3). Ideally, the impact evaluation of the ISA is to compare outcomes between the situation "with" ISA and the situation "without" ISA for a given municipality or district. Unfortunately, it is not possible to observe the situation "with" ISA and the situation 'without' ISA for any given municipality and district in the same time. In practice, the effects of ISA is estimated by comparing outcomes between municipalities (districts) receiving ISA and those not receiving ISA before and after the intervention. To argue that the identified effects of ISA through the evaluation exercise described in this section to be causal, we implicitly assume that all the other nation-wide activities under Activity I and Activity II have the same effects across municipalities and districts and the change in outcomes between the treatment and control municipalities/districts remain constant in absence of the intervention. We will come back to the evaluation method later.

For better exposition of the intended goal of this evaluation effort, we present in table 3 the impact pathway of the ISA, focusing on linking the activities to short-term outputs and then consequently to the outcomes the intervention aims to achieve. As indicated in table 3, the ISA (i.e., investments made to upgrade the municipal or district land administration systems) would lead to a number of outputs at the municipal or district level (e.g., increased number of clients aware of the land law, increased number of Cadastral officers trained, upgraded facilities). The ultimate objective of these activities is to increase efficiency and effectiveness of the municipal and district land administration offices which is explicitly measured by a number of outcome indicators the key of which include (1) reduced processing time to obtain DUAT, (2) reduced cost of obtaining DUAT; (3) increased number of DUATs demanded, (4) increased number of land transactions, and (5) reduced incidence of land disputes.

Table 3: Impact pathway for the Institutional Strengthening Activities (ISA) of the Land Project

ISA Activities	Outputs	Outcomes	Impact indicators
• Professional development and training, upgrading of facilities in 8 municipalities and 12 districts in 4 northern provinces • Assistance to cadastral development in 8 selected Municipalities	• Outreach of Land Law and use rights conducted • Comprehensive approach to professional development and training implemented • Improvement of the Land Administration System (LIMS) of the 8 municipalities and 12 districts	• Increased awareness of Land Law • Increased number of Cadastral officers trained • Upgraded facilities and IT equipment • Improved/more effective operational procedures	• Reduced processing time to obtain a DUAT; • Reduced cost of obtaining a DUAT; • Increased demand for DUATs; • Increased number of land transfers; • Reduced number of land disputes.

3.2. Geographic coverage of the IE and identifying comparison group

The ISA were implemented in all the 8 municipalities and 12 districts as listed in Tables 1 & 2. Due to the fact that the municipalities and districts serve very different clients who demand DUATs for different types of land (urban land versus rural land), it makes more sense to evaluate the ISA in municipalities and in districts separately. Based on the discussion in the previous section, one of the critical steps for a rigorous impact evaluation is to identify reliable counterfactual group (control group). According to Tables 1 and 2, the 8 municipalities and 12 districts were chosen according to a wide variety of selection criteria. It is important to identify control municipalities/districts that face similar issues as the treatment municipalities/districts. We classify all the eight selected treatment municipalities into seven different groups (A, B, C, D, E, F, G), and the twelve districts into seven different groups (I, II, III, IV, V, VI, VII) as well according to the selection criteria and similarity in geographic location and the condition of local land administration system prior to intervention. The classification of the municipalities and districts is indicated in the last column of Tables 4 and 5.

Next for the treatment municipalities/districts in each category, at least one control municipality/ district is selected that match both the selection criteria and the local conditions before 2009 (the originally planned project implementation start time). Specifically, the local condition used for the matching includes the number of staff in cadastral service, size of the cadastral unit, average number of years of experience of cadastral staff members, average number of previous trainings conducted, quality of equipment in the cadastral office, number of DUATs applications processed within 90 days per year or month, quality of facilities (access to electricity, number of survey equipment by type). To further improve the reliability of the control group, we will also select the control municipality/district from those that had also applied for the Project but were not selected. As a result, 7 control municipalities and 7 districts were selected to match with the 7 categories of the treatment municipalities and treatment districts, respectively (column 3, Table 4 and Table 5).

Table 4: Selected treatment and control municipalities

Province	**Treatment Municipality (N=9)**	**Control Municipality (N=7)**	**Group**
Nampula	Monapo	Angoche	E
	Nampula city	Nacala-Porto	D
Cabo Delgado	Pemba	Mueda	C
	Mocimboa da Praia	Chiure	B
Niassa	Lichinga,	Marrupa	A
	Cuamba		
	*Metangula**		
Zambezia	Quelimane	Alto Molocue	G
	Mocuba	Gurue	F

***Metangula* is a small municipality near Lake Niassa that was added to the list of municipalities where land regularization occurred. This was done because Metangula is small and HTSPE thought it could complete the cadaster 100% in the municipality. It was intended toinclude at least 1 municipality to complete and establish its cadaster so that all records could be accurate and reliable.**

Table 5: Selected control districts

Province	Treatment	Control (N=7)	Group
Zambézia	Nicoadala		I
	Morrumbala	Alto Molocue	II
	Mocuba		
Nampula	Malema	Nampula	III
	Monapo		
	Moma	Nacala	VI
Cabo Delgado	Mocimboa da	Pemba	IV
	Montepuez		
	Mecufi	Palma	V
Niassa	Majune	Cuamba	I
	Lichinga		
	Metangula	Mandimba	VII

3.3. Identification issues

The effects of institutional strengthening can be identified using DID approach as described in section 2.2. There will be a minor modification of equation (2) in the econometric specification. Specifically, equation (2) is a standard DID model for data from two time periods (i.e., before intervention and after intervention periods). The multiple-year administrative land data prior to the intervention that were assembled by the MSU team served as baseline data, and similar administrative land data for several years after the intervention will be collected as endline data. The data issues are discussed in the next section. From an econometrics point of view, it is always better to have data from more years. To account for the fact that data for multiple years before the intervention and multiple years after intervention may be available, we need to modify equation (2) as the following:

$$Y_{ij} = \alpha + \beta T_j + \boldsymbol{\gamma t} + \boldsymbol{\delta}(T_{ij}*\boldsymbol{t}) + \boldsymbol{\rho X_i} + \varepsilon_{ij} \qquad (3),$$

Where, subscript i stands for a specific district or municipality, j for a specific output, T_i is the treatment dummy variable distinguishing treatment district or municipality (=1) from control districts/municipalities (=0), **t** is a vector of time dummies for different years, $\boldsymbol{X_i}$ is a vector of other district/municipality or parcel level control variables, β captures the regional difference between the treatment and control districts/municipalities, **γ** captures the common time trend effects over time in all districts/municipalities, and **δ** is the vector of parameters of interest, measuring the impact of institutional strengthening program on outcomes of interest (e.g , transaction time, number of transactions, etc.), and ε_{il} is the idiosyncratic error that is assumed to be normally distributed with mean zero and unity variance. The key difference between (3) and (2) is that the time dummy variable (t) in equation (2) is now replaced by a vector of time dummies ***t***. It is expected that the coefficients of the time dummies for the years after the intervention should be significant and have the expected sign (e.g., negative in terms of time reduction in the process of DUAT issuance and positive in terms of number of DUATs issued, etc.), and we expect the coefficients of the time dummies for the period before the intervention to be not significantly different from zero (the parallel trend assumption).

3.4. Data collection

Based on the discussion above, identifying the program effects of the institutional strengthening activities will depend on administrative data from the treatment and control municipalities or districts before and after the program implementation. **The first important question related to data collection is how many years of data are needed for the analysis?** In general, we would like to have as many years as possible based on availability of the administrative data. An MSU team member visited Mozambique in the summer of 2013 and collected the baseline data for the study. He collected administrative land record data as early as in 1980 with the assistance of each land administrative office. The main purpose of collecting data from the pre-reform years was to allow the test for pre-trend difference between the treatment and control municipalities/districts. While it is not necessary to have data from distant past to test/control for the pre-program trend, there is no harm to have as many years as possible in the analysis. In other words, there is no need to restrict the use of historical data.

On the other hand, more care is needed when it comes to using data from the post-reform periods. Specifically, we cannot use the data from the years during and immediately after the implementation of the activities because of the contamination effects of the program (i.e., artificially speedy DUAT process and low cost of DUAT issuance), it is therefore, important to exclude data from these years. While there is no standard rule as to how many years should be excluded, we think it is reasonable to assume that the contamination effects is unlikely to be significant after 2 years from the completion of the intervention. During the visit to collect the baseline data for this project, the MSU team determined that the implementation of the institutional strengthening activities took place in 2012 in almost all the municipalities or districts, so the administrative land record data from 2012 to 2014 should be excluded. As for when the end-line data collection after these few years of buffer period should be implemented, one critical rule is that data collection should be conducted before any of the control areas starts to receive similar program activities. We recommend the end-line survey to be conducted in 2018/2019 so there will be about 4-5 years useful, post-reform data for analysis.

The data used for this study are from two sources: administrative land record data on urban land from the municipal cadastral office (for the municipality evaluation) and administrative land record data on rural land from district SPGC offices. While the administrative land record data from the district SPGC offices are available in digital form, the data from municipal cadastral offices are in paper form. For the baseline data collection, it took little effort to copy the digitized data from district SPGC offices, but a lot of time and effort was spent by the MSU team on converting the paper form land administrative data from the municipal cadastral offices into digital form.

The second important question is which variables in the land administrative data set should be used? While ideally we would like to have parcel level characteristics, and detailed characteristics on the owners of parcel who apply for DUAT, we are facing the reality of a limited number of variables available in the administrative land record data. To meet our objectives to evaluate the impacts of the institutional strengthening activities on three main outcome indicators, namely, the total number of DUATs issued, the time taken to process a DUAT (from application to issuance), and the number of land transfers, we collect administrative record data on: (1) time when the application was filed; (2) time when an DUAT was issued; (3) time when each land

transfer occurred. Besides these key variables, we also have data on the following variables that should be included in equation (3) as control variables:[6]

(a) Parcel size;

(b) Main use of a parcel from municipality data: residence, commerce, industry, social/religion, public services, residence & commerce;

(c) Main use of a parcel from SPGC data: Agricultural production (annual & perennial crops), forest plantations, livestock production (cattle & others), public services, commerce & industry, residence, tourism, social & religion, crop-livestock production, community;

(d) Gender of the DUAT holder

3.5. Some useful lessons from the baseline data

An earlier attempt to evaluate the activity 2 in districts using the Agricultural Survey Project (*Trabalho de Inquérito Agricola* – TIA) data was ineffective due to the representativeness of MCC intervention sites in the sample of TIA. Specifically, TIA was representative at the provincial level so its sample is not appropriate to address issues at either the district or manicipal level. The report produced and submitted to MCC only provided indicative figures as it compared treatment and control districts (municipalities).

The main challenge to this evaluation has to do with the data quality (completeness, reliability, validity). The number of DUATs at baseline, pre-project available to the evaluators is likely to be small due to poor record storage/keeping which can compromise the reliability of the estimates (for e.g. in Mocimboa da Praia a significant number of old DUATs were lost due to floods, in Marrupa due to not having a copy machine original copies were given to the applicants and no alternative copies were kept in the cadastral office).With these challenges, we advise that MCC be aware of the potential limitations of this evaluation exercise. We recommend using the entire population in the treatment sites on multiple years to obtain as many valid records as possible for the evaluation.

Second, the success of the whole evaluation process depends on the collaboration and access to quality data from the local land administrative units. Therefore, we recommend that municipal cadastral offices and district SPGC offices improve record-keeping system to keep transaction records in a systematic and easily accessible manner.

Finally, we are not able to evaluate the impacts of this program on cost savings associated with DUAT process or number of land disputes/conflicts because the administrative data have no information on these variables. While collecting information on these important variables will not help us much in this impact evaluation due to the lack of this information in the baseline, it will be valuable to compare whether the cost and disputes are indeed lower in municipalities or districts that were affected by the institutional strengthening activities, especially if the assumption that the difference between the treatment and control groups prior to the intervention is similar.

[6] Although we are also interested in the monetary cost, such information is not available.

4. Impact Evaluation Design of the Urban Hotspot Activities Under Activity III of the Land Project

4.1. Overview of the activity being evaluated

A list of the eight municipalities and the selection criteria they meet for Land Project activities is given in Table 1. The priority areas (or *bairros*) identified for site specific activities within these eight municipalities are the smallest unit of project interventions of the Land Project in urban areas. As such, the impacts to be observed at the beneficiary level in these priority areas (or *bairros*) will be a cumulative sum of all the three pillar activities of the Land Project (i.e., policy, capacity building and site specific activities).

Project activities will be implemented with technical assistance from service providers such as CENACARTA, and the implementing partner (HTSPE) and cover the following major activities:

a) The satellite mapping and inventory exercise
b) Piloting a sound approach to area-wide registration of land rights

Activities under 'a' are generic at the level of all selected 8 municipalities (i.e., cover all the bairros within the selected municipalities). However, activities under 'b' will be implemented only in selected priority bairros to address some hotspot issues related to expansion, requalification and regularization (Annex 1). The purpose of this area-specific interventions in priority bairros is to pilot a sound approach to area-wide registration of land rights to individuals.

4.2. Geographic coverage of the IE

Given the time and resource constraint, it was not feasible to implement a rigorous IE study for all eight municipalities. The two municipal areas in the Province of Nampula--Monapo Vila and Nampula city were selected for this rigorous impact evaluation of urban hotspot activities based on discussions with MCC/MCA and taking into account the following criteria:

- Large numbers of bairros facing the same hotspot issues (i.e., expansion, requalification and regularization issues – see Annex 1) in a given municipality.
- Ability to identify comparison bairros to estimate the effects of the intervention in a rigorous and robust manner.
- Indication that project interventions in hotspot areas are planned earlier in MCA's 5-year implementation plan (to ensure enough time to observe outcomes and impacts).

The focus of the impact evaluation design described in this document is to assess the impact of the interventions targeted in priority bairros within Nampula city and Monapo Vila that have been identified under the hotspot issue of "requalification / regularization."[7] The geographic coverage includes four priority bairros in Nampula city and six priority bairros in Monapo vila (Table 6). These were selected and prioritized by the municipalities based on some set criteria and were outside the control/influence of the impact evaluation team.

[7] In other words, the hotspot issue of 'expansion' is not included in this evaluation design.

Table 6. Project intervention bairros for "hotspot" site specific activities under Pillar III

Nampula City	Monapo Vila
Muatala Muhala – Sede Mutauanha Namutequeliua	Mucaca Mecutane Topelane Moajem Boa Viajem Metropime

4.3. Research questions to be addressed by the IE

Initiatives to strengthen the property rights system as envisaged in Nampula city and Monapo Vila are generally designed to result in clearly defined rights that are enforceable, transferable, and of appropriate duration and scope. An improved system should lower land-transaction costs, lower the risk of expropriation or conflict, and increase tenure security. In the medium or longer term, the system should contribute to more efficient land uses due to improved productivity, increased investment, and the development of land markets. More productive land should result in higher asset/land values and higher incomes for property owners. Over time, as land and financial markets develop, formal land rights can also be used as collateral for loans.

Empirical studies suggest that impacts of land tenure projects vary considerably from country to country, depending on market development, financial institutions, legal frameworks, and beneficiary income. Land tenure reform has demonstrated impacts for economic growth that reaches the poor, but can have socially differentiated impacts that need to be measured and monitored. The purpose of the rigorous IE design for the two urban hotspot areas is to precisely measure and monitor these impacts. The key research questions guiding our design of the evaluation for urban hotspot activities in Nampula City and Monapo Vila are to evaluate the extent to which there is evidence of change in indicators of outcomes and impacts listed in Table 7.

Table 7. Impact pathway for area-specific intervention in priority bairros (i.e., to address the hotspot issue related to requalification/regularization)

Activities	Outcome	Impact indicators
• Digitized base maps for "priority areas" • Demarcated plots • Issuing DUAT for the plots	• Make the process simple, cost-effective and faster • Increased security of tenure	• Reduced incidents of conflicts • Increased new commercial enterprises and activities • Increased level of investments on land parcels • More effective/productive land uses • Increased off-farm opportunities (labor mobility) • Increased access to formal credit (i.e., collateral effect)

4.4. Identifying Comparison Communities

There are two things needed to implement the DID IE design:
1. Identification of treatment and comparison sites, and
2. Data collection from both treatment and comparison sites before and after intervention.

The prioritized bairros listed in Table 6 are the potential pool of treatment sites for this IE. The units of impact observation will be households. Thus, households within the boundary of these listed bairros serve as the treatment group. If the time line for implementing the interventions in prioritized hotspots was such that project implementer could have staggered the implementation across these bairros over time, ideally, we could have implemented a 'pipeline' design whereby the order of project intervention across prioritized bairros could have been randomized. In that scenario, bairros randomly assigned to receive intervention in the first year could have served as treatment and bairros randomly assigned to receive the intervention in year 5 could have served as control. However, based on the discussions with municipal staff and project implementing partners, it is clear that a pipeline design is not feasible for these two selected municipalities. The reason is that the intervention bairros have been already prioritized from among a pool of all potential bairros in the municipality, and in the case of Monapo they have been assigned a priority order.

Given this reality, we are using the following strategy in each of these two municipal areas to ensure we have sufficient number of comparison households to implement the DiD design.

For Nampula, the strategy is to select an additional bairro (Muahivire) that is facing the same hotspot issue but is not in the priority list. Baseline data will be collected from all five bairros—the four priority bairros and one non-intervention bairro. The plan is to over sample households in this non-intervention bairro (Muahivire) as it will serve as a comparison site for the IE. Any bairro that does not receive the intervention by Year 5 (before the follow-up survey), will also serve as an additional comparison site for the IE.[8]

For Monapo, we are following a similar strategy but the numbers are different. We have selected the following five bairros (which are all peri-urban) to serve as comparison bairros.[9] The bairros not selected as part of this IE design from Monapo Vila were all rural bairros.

Mulotine
Nachicuva
Naheruque
Micolene
Nova Cuamba

In addition, if any of the seven priority bairros listed in Table 6 do not receive intervention before the follow-up survey planned in year 5, then that bairro will also serve as a comparison bairro.

---

[8] Given the large size of each bairro in Nampula, it is likely that it may take more time to complete all the intervention activities in four bairros. If the interventions are undertaken in a sequence and it takes an average one year to complete one bairro, then this scenario is potentially possible.

[9] Ideally, we would have preferred an evaluation design that had a mix of urban and peri urban bairros in both the treatment and control sites. However, since the municipality has already selected priority bairros (which are all urban bairros) and the order in which they will be treated, we are left with only peri-urban bairros for control group.

Thus, the IE plan consists of conducting baseline and follow-up surveys in five bairros in Nampula (4 priority + 1 extra) and in 11 bairros in Monapo Vila (6 priority + 5 extra). All the bairros to be surveyed are listed in Table 8 along with some key characteristics based on 2007 census data.

Table 8. List of selected bairros for the impact evaluation design in Nampula ciudad and Monapo Vila: Main characteristic features[\a]

Selected Bairros for the IE design	Urban (U) / Peri Urban (PU) / Rural (R)	Priority (order) Given by municip-ality	Potential treatment or control sites	No. of total Enumer-ation Areas (EAs)	No. of HHs	Popul-ation (N)	% of hhs with farm income	% of HHs with TV	% of HHs with female head
Nampula:									
Muatala	U	Yes	Both	77	9731	45231	23.57	26	22
Muhala – Sede	U	Yes	Both	76	11380	59618	20.91	37	21
Mutauanha	U	Yes	Both	72	13438	62976	29.04	25	18
Namutequeliua	U	Yes	Both	51	9405	45154	26.12	20	22
Muahivire	U	No	Extra control	78	11052	49763	32.80	23	20
Total					55,006				
Monapo:									
Mucaca	PU	Yes (3)	Trtmt	9	1108	4392	39.98	6	25
Mecutane	U	Yes (4)	Trtmt	8	743	3549	55.45	14	31
Topelane	U	Yes (7)	Trtmt	7	676	2785	43.05	15	27
Moajem	U	Yes (8)	Trtmt	5	489	2395	47.44	17	26
Boa Viagem	U	Yes (9)	Trtmt	5	537	2486	44.69	20	21
Metoprime	U	Yes (10)	Trtmt	4	386	1773	67.36	14	27
Mulotine	PU	No	Cntrl	6	610	2925	54.59	17	34
Nachicuva	PU	No	Cntrl	21	2008	8142	47.06	6	25
Naheruque	PU	No	Cntrl	6	508	2164	27.36	5	23
Micolene	PU	No	Cntrl	8	477	2041	57.65	9	30
Nova Cuamba	PU	No	Cntrl	9	1355	5576	43.10	6	29
Total					8,897				

\a Characteristics and statistical data are based on the sample frame and results of the 2007 census. (Post-script: Data collection efforts prior to baseline survey to finalize the sampling frame and to identify a list of eligible households indicate that the sampling frame has deteriorated substantially over the last 4 years, especially in Monapo Vila, with a concomitant decline in population.This may be a consequence of migration of households from small urban areas to larger cities or movement to locations outside the frame with new settlements near roads, etc.

4.5. Sample size and Power Calculation

The power of the design is the probability that, for a given effect size and a given statistical significance level, we will be able to reject the hypothesis of zero effect. Sample sizes, as well as other design choices, will affect the power of an experiment.

To estimate the total sample size for this IE design, we treat Nampula city and Monapo Vila as two independent evaluations, but both addressing the same impact questions for similar interventions. For each of these two urban areas, we follow the steps described below (and elaborated in Table 8) to estimate the total sample size.

In step 1, we applied the power calculation based on a simple random sampling method using the formula in equation (4) to estimate the minimum required sample size for Nampula city and Monapo Vila based on the following parameter values: a power (k) of 80% (i.e., t_{1-k}=0.84), a significance level (α) of 0.10 ($t_{\alpha/2}$=1.65), and portion of subjects allocated to treatment group (P=0.5), and a standardized minimum detectable effect size (MDE), m=(MDE/σ) of 0.25.

$$n = \left[\frac{(t_{1-k} + t_{\alpha/2})^2}{m^2 * P(1-P)} \right] \qquad (4)$$

Equation (4) is basically the same equation (7) in Duflo et al. (2007). The only difference is that we use to solve for sample size rather than for MDE and the *m* in equation (4) is the standardized MDE (i.e., minimum detectatble effect size divided by standard deviation).

The estimated minimum sample size based on this formula and the given parameter values noted above came to 397 for each city (Table 9). Table 10 shows how the sample size would change under different parameter values to achieve the power of 80%. For example, the number would change to 413 if we change P to 0.6. Alternatively, with P=0.5, the sample size estimate is 501 if we change α to 0.05. The corresponding number of observations for α=0.10 (or 0.05) would further increase to 620 (or 780) if we set *m* at 0.2 instead of 0.25.

In reality, as a means of saving money, the simple random sampling is rarely used because it requires the researcher to sample across all geographic areas within the domain. Thus, cluster sampling is more common than a simple random sampling approach. In this IE design, we also plan to follow this practical approach and sample households from a sub-set of enumeration areas (EAs) within a given bairro. This cost saving measure, however, does reduce the confidence level of the estimates for a given sample size. This loss of effectiveness by the use of cluster sampling, instead of simple random sampling (SRS), is the **design effect,** defined as the ratio of the actual variance under the sampling method actually used, to the variance computed under the assumption of simple random sampling.

Table 9: Steps used in estimating the sample size for the IE design

Steps	Parameters	Nampula	Monapo Vila
1: Apply "Simple Random Sampling" method	Power (k) of 80%	80%	80%
	Significance level (α)	0.10	0.10
	Portion of subjects allocated to treatment group (P)	0.5	0.5
	Standardized minimum detectable effect size (MDE), m=(MDE/σ)	0.25	0.25
	Estimate of minimum sample size (SRS)	**397**	**397**
2: Adjust for the design effect	Design effect (DEFF)	2.0	2.0
	Effective sample size = SRS * DEFF	**794**	**794**

3: Adjust for attrition from baseline to follow-up survey	Attrition factor	**13%**	**11%**
	Adjusted sample size = Effective sample size * (1+ attrition rate)	897	881
	Sample Size (Rounded off)	**900**	**880**

Table 10: Sample size required to achieve the power of 80% under different parameter values

		(MDE/σ)=0.25		(MDE/σ)=0.20	
P	1-P	α=0.10	α=0.05	α=0.10	α=0.05
0.50	0.50	**397**	501	620	780
0.60	0.40	413	523	681	861
0.65	0.35	436	551	646	816
0.70	0.30	472	597	738	933

In general, using a cluster sample generally requires either a larger sample size than a simple random sampling or using a wider confidence interval. The design effect is used to determine how much larger the sample size or confidence interval needs to be. The main components of the design effect are the intraclass correlation, and the cluster sample sizes. Given the fact that we are potentially interested in many outcome variables in this IE design and the data requirement at the EA level from previous surveys to estimate the intraclass correlations for all the outcome indicators, which were not available to us, we used a simplistic approach of assuming the design effect to be 2.0. Most studies in the literature report a design effect in the range of 1 to 3[10] (Shackman 2001); so this assumption of a design effect = 2 is not unrealistic.

In second step, the estimated sample size from SRS was multiplied by the design effect (2.0) to get an effective sample size (Table 9). However, given the potential attrition rate for the longitudinal survey, in step 3 we increased the sample size for both the urban areas in the baseline survey by a factor of 13% for Nampula (which is more urbanized) and 11% for Monapo (which is peri-urban and more rural). The end result of all the three steps is an estimated total sample size of 900 households for Nampula city and 880 households for Monapo Vila (Table 9). These are the target sample size for the IE design in the two urban hotspot priority areas.

4.6. Sampling Method

Once the sample size is determined as described above, the actual selection of the sample of households is done as described below. Depending on the number of enumeration areas (EAs) (which are the primary sampling units in the context of Mozambique sampling frame) and the total sample size targeted for the survey (described above), the sampling method for this IE design will follow the following one- or two-stage sampling design (Table 11). The sampling frame for the purpose of this IE is defined as "households that have land in the given municipality."

<u>Two-stage sampling design</u>

[10] Some studies also report design effects less than 1 and more than 3.

For all the treatment and control bairros in Nampula city the following two-stage sampling design will be used: In stage one, we propose to randomly select 10 EAs for the 4 prioritized hotspot bairros and 20 EAs for the additional control bairro. This sample of 10 or 20 EAs will be selected within each bairro systematically with probability proportional to size, where measure of size is the number of households based on data from the Population Census of 2007.

In stage two, 15 households in each EA will be randomly selected which will give a total sample of 900 households in Nampula City (across both potential treatment and control bairros). The random selection of 15 households in each EA will be based on the 'table of random numbers' generated for potential size of the EAs ranging from 40 households to 450 households.

For one of the control bairros in Monapo Vila--Nachicuva, which has 21 EAs, this same two stage sampling process will be used, but with a different number of target EAs and households. In the first stage, 5 EAs will be selected in Nachicuva with probability proportional to size of the EAs (based on data from 2007 population census). In stage 2, 16 households will be selected from Nachicuva based on the 'table of random numbers' as described for Nampula city.

<u>One-stage sampling design</u>

For all the remaining 10 small Bairros in Monapo Vila that have less than 10 EAs we plan to adopt a one-stage sampling design to simplify the estimation procedures. The plan is to select 80 households from a list of all the households ordered by EAs in a given bairro based on a method called 'systematic sampling with a random start'. With a one-stage random systematic selection of households from the list of households ordered by EA, the number of sample households selected in each EA would be proportional to the size of the EA, thus simplifying the estimation procedures.

Table 11. Number of EAs and households in each selected Bairros included in the baseline survey in Nampula ciudad and Monapo Vila

			Stage 1: Selection of EAs		**Stage 2: Selection of HHs**
Selected Bairros for the IE design	Potential treatment or control sites	No. of total Enumeration Areas (EAs)	No. of EAs to be selected	No. of HHs in selected EAs that meet the sample frame criterion	No. of HHs to be selected for baseline survey
Nampula:					
Muatala	Both	77	10	453	150
Muhala – Sede	Both	76	10	638	150
Mutauanha	Both	72	10	483	150
Namutequeliua	Both	51	10	1034	150
Muahivire	Extra control	78	20	1391	300
Total					900
Monapo:					
Mucaca	Trtmt	9	All \a	563	80
Mecutane	Trtmt	8	All	305	80
Topelane	Trtmt	7	All	383	80
Moajem	Trtmt	5	All	357	80
Boa Viagem	Trtmt	5	All	186	80
Metoprime	Trtmt	4	All	154	80
Mulotine	Cntrl	6	All	310	80
Nachicuva	Cntrl	21	5	189	80

Naheruque	Cntrl	6	All	221	80
Micolene	Cntrl	8	All	239	80
Nova Cuamba	Cntrl	9	All	695	80
Total					880

\a For bairros that had less than 10 EAs, "All" the EAs are included in a one-stage systematic selection of households.

4.7. Data collection

The evaluation will use household level surveys that will include interviewing the head of the household based on a detailed instrument. The questionnaire includes more than 25 sections encompassing modules on:

1. Household characteristics (demographic information by each member of the HH)
2. Employment and sources of any other cash transfers
3. Identification and list of all the parcels
4. Land conflicts
5. Rights to the land and perceptions of the risk
6. Parcels rented out, rented in
7. Characteristics of parcels
8. Investments on land
9. Perceptions about the DUAT, renting land and the land law
10. Relative space occupied by crops in the plot
11. Production and sales of basic food crops, cash crops, vegetables, fruits, nuts, etc.
12. Agricultural practices
13. Ownership of Assets
14. Monthly expenditures
15. Credit in the last 12 months
16. Livestock and sub-products produced and sold in the last 12 months
17. Consumption

The survey has detailed sections for each of the outcomes to be evaluated, both intermediate and final outcomes. In addition, each of the survey households will be geo-referenced. If the head of the household is not present at the time of the first visit, enumerators will attempt to make an appointment and return again to interview the appropriate person, provided that this return visit is possible within the time that the survey team will be in the area. In households that are male-headed with a spouse present, the spouse will be the respondent for the livestock and food consumption modules. The survey is designed to take between 1 and 1 ½ hours.

<u>Survey Calendar</u>

The baseline survey was implemented from October to December 2010. There experienced some delay in the implementation of the projects. As a result, DUATs were not issued until late - 2012/2013. Taking into account this long delay in the project implementation and the fact that it takes several years for the real impact of a DUAT to realize, we recommend that the follow-up survey is scheduled in 2017 or 2018.

4.8. Assessing the original design after project implementation (as of May 2016)

4.8.1. Validity of the original design for Monapo vila

The implementation of Activity III in Monapo vila did not exactly follow the original implementation plan. In fact, all five Bairros in the Monapo Vila that were selected as control Bairros in the original design were also treated by the end of the compact, which leaves no single control Bairro for the evaluation. This was further confirmed during the recent field trip in May 2016 **As a result, there will be no valid impact evaluation of Activity III in Monapo vila.**

4.8.2. Validity of the original design for the Nampula City

The implementation of Activity III in Nampula city also did not follow its original plan, which affects the validity of the original evaluation design. There are two specific issues that deserves our attention. The first issue is related to the slow progress of the project implementation. For example, only 8% of the households in the treatment Bairros received treatment in the first year after the implementation began. This issue was reconfirmed by the field trip in May 2016. By the end of the compact (and by May 2016), only 295 households out of the 560 households that were intended to be treated were actually treated. The second issue is that the intervention was implemented at the Unidade Comunal (UC) level rather than at the Bairro level. In other words, the actual implementation moved from one UC to another only after the work in the first UC was completed. Because UC is a sub-community within a Bairro, the change from Bairro to UC as the intervention unit would lead to a significant increase in number of clusters (from the original **5** bairros to **29** UCs), which should help increase the statistical power. Both the slow implementation progress and the change of intervention unit have important implication on the validity of our original design.

To help assess whether we can still use the sample and baseline survey from the original design to conduct a rigorous impact evaluation of the Activity III in Nampula city, we recalculate the statistical power based on the most updated information collected from the field trip in May 2016. In the remaining section, we present the recalculation of the statistical power.

We again follow a formula used by Purdon (2002) and Duflo et al. (2007) to calculate the statistical power of a group-based program design. The most relevant parameter in the context of impact evaluation is the minimum detectable effect size (MDE), that is the smallest program effect that can be detected at a certain level of confidence based on a given set of statistical parameters (i.e., a given power (k), a given level of statistical significance (α), sample size and the proportion of sample allocated between the treatment and control groups. Specifically, the formula to obtain MDE or standardized MDE (MDE/σ) given all the necessary parameters is:

$$MDE = \frac{M_{J-2}}{\sqrt{P(1-P)J}}\sqrt{\rho + \frac{1-\rho}{n}}\ \sigma, \quad \text{or} \quad \left(\frac{MDE}{\sigma}\right) = \frac{M_{J-2}}{\sqrt{P(1-P)J}}\sqrt{\rho + \frac{1-\rho}{n}} \qquad (5)$$

Where, M_{J-2} is the minimum detectable effect multiplier the magnitude of which is determined by the level of significance, power and the number of groups included in the intervention (Bairros or EAs in our case), P is the proportion of sample allocated between the treatment and control groups, ρ is the intracluster correlation coefficient, J is the total number of groups included in the experiment, and finally n is the number of households in each group. The detail explanation of each of the parameters and the derivation of the formula can be found in Duflo et al. (2007).

Table 12: Current situation in the municipality of Nampula, 2016

Bairro	Unidade Comunal (UC)	Initial classification (2010)			HTSPE intervention			MSU-HTSPE matching (2013)				Not identified (N)	HTSPE intervention?	
		Treat.	Control	Total	Outside	Within	Total	Unmatched	Partial	Matched	Total		Yes/No	Coverage
Muhala-Sede	25 DE JUNHO	14	0	14	1	13	14	6	0	7	13	1	YES	ALL
	7 DE ABRIL	75	0	75	74	1	75	1	0	0	1	1	NO	
	EDUARDO MONDLANE	28	0	28	28	0	28					0	NO	
	JOSINA MACHEL	15	0	15	15	0	15					0	NO	
	PAULO SAMUEL KANKOMBA	15	0	15	15	0	15					0	NO	
Namuteque-liua	NAMALATE	15	0	15	15	0	15					0	NO	
	AMILCAR CABRAL	15	0	15	0	15	15	1	4	10	15	0	YES	ALL
	MUTOMOTE	60	0	60	0	60	60	15	15	30	60	1	YES	ALL
	MIRIAN NGUABI	60	0	60	3	57	60	11	13	33	57	2	YES	ALL
Muahivire	ELIPISSE	30	0	30	4	26	30	9	4	13	26	1	YES	ALL
	GORONGOZA	15	0	15	15	0	15					0	NO	
	MUETAZE	60	0	60	25	35	60	9	7	19	35	2	YES	PARTIAL
	RENO	14	0	14	2	12	14	6	1	5	12	0	YES	ALL
	MUACOTHAIA	15	0	15	5	10	15	3	2	5	10	1	YES	PARTIAL
	MUCUACHE	30	0	30	30	0	30					0	NO	
	MUENGANE	45	0	45	4	41	45	14	10	17	41	0	YES	ALL
	MUTOTOPE	30	0	30	10	20	30	6	4	10	20	3 (2+1)	YES	ALL
	NAMUATO	45	0	45	16	29	45	10	7	12	29	0	YES	PARTIAL
	NANUCO	15	0	15	14	1	15	1	0	0	1	1	YES	PARTIAL
LOST SAMPLE (unidentified in 2016)		**13**	**0**	**13**	**0**	**13**	**13**	**9**	**4**	**0**	**13**	**13**		
LOST SAMPLE (transactions)		**22**	**0**	**22**	**0**	**22**	**22**	**15**	**7**	**0**	**22**	**0**	**25 households lost**	
Mutauanha	25 DE SETEMBRO	0	15	15	15	0	15						NO	
	7 DE SETEMBRO	0	14	14	14	0	14						NO	
	EDUARDO MONDLANE	0	29	29	29	0	29						NO	
	MUTHITA	0	15	15	15	0	15						NO	
	PILOTO	0	67	67	67	0	67						NO	
Muatala	COSSOLE	0	30	30	30	0	30						NO	
	MINICANE	0	15	15	15	0	15						NO	
	MURALENE	0	15	15	15	0	15						NO	
	NAMAVO	0	70	70	70	0	70						NO	
	NAPALA	0	15	15	15	0	15						NO	
Total (original)		*596*	*285*	*881*	*561*	*320*	*881*	*92*	*67*	*161*	*320*	--		
TOTAL (final)		**571**	**285**	**856**	**561**	**295**	**856**	**68**	**56**	**161**	**295**	**13**		

To obtain (MDE/σ), we need to assign values for all the parameters on the right hand side of equation (5). Again, given the conventional level of significance and the statistical power (i.e., α=0.05 and β=0.8), and J=29 (the total number of intervention units), the corresponding M_{J-2} in equation (5) is 2.85. Knowing that the total sample size for the treatment UCs and the total sample size for all the control UCs, we can easily obtain the share of households in the treatment UCs (i.e., P=295/856=0.37). We also know the average sample size in each UC is approximately N=30. Finally based on the baseline survey data, we obtain the average intracluster correlation coefficients of the key variables, ρ=0.014.

Plugging the above parameter values into equation (5), we will obtain (MDE/σ)=0.241, which is between the conventional "small" (0.2) and "medium" (0.5) according to Duflo et al. (2007). Actually it is barely above the standard of "small". So it is a reasonable detectable effect size.

4.8.3. Justification of MDEs for non-randomization and additional co-variates

The MDEs presented above are based on randomized experimental design. Any deviation from the randomized experimental design will need adjustment in power calculation. Since the bairros were selected into treatment and control groups according to their prioritized rank, the original design was not an experimental design. To be more precise, the design is a difference-in-difference quasi-experimental design. There is no equivalent formula to calculate the power of any non-experimental design. According to Purdon (2002), in practice the sample size (or power calculation) for a quasi-experimental design is based on the same formula for randomized design (equation 1). The difference is that a quasi-experimental design needs to have somewhat larger number of sample size, especially in the control group to allow for the fact that the data from the comparison group may need to be 'adjusted' (either by weighting, matching, or by statistical modeling) to make it comparable with the treatment group (Purdon 2002). Considering the relatively low MDEs, we will still get reasonably acceptable MDEs even if adjust the sample to a much smaller size (e.g., reducing 20% for all bairros). For example, the sample size per UC in would be reduced from 30 to 20. The adjusted MDE is now 0.35, again between "small" and "medium". In the meantime, adding additional covariates in the regression model is likely to further offset the negative effect caused by the non-randomized design. So we are confident that the current situation would still allow us to conduct a rigorous impact evaluation in Nampula.

4.9. Implications for next steps on evaluation of Activity III in Nampula

The revised power calculation (small MDE) confirms that a rigorous impact evaluation of Activity III in Nampula city is possible despite the deviations of the project implementation from its original plan. It is important that there is no further change in the treatment status of the sample households. In other words, there should be no other land titling project in the treatment UCs and the control UCs from now to the time before the endline survey is implemented. We recommend the endline survey to be conducted in 2017/2018 to allow sufficient time for the real impact to realize and to avoid further contamination (if it takes too long).

5. Impact Evaluation Design of the Rural Hotspot Activities Under Activity III of the Land Project

5.1. Overview of the activity being evaluated

In the context of rural areas, the rural hotspot interventions planned in selected districts include support to the formalization of land use rights, the systematic planning of development areas and the parceling of land plots for subsequent attribution. It also includes the development of civic education materials and communication initiatives at local level (including seminars, workshops and public hearings) and support to local authorities in providing the public with up-to-date information on the land use and land tenure status of particular areas. The site-specific interventions in priority Aldeias that are subject of this impact evaluation include following activities that are implemented with technical assistance from service providers such as CENACARTA, and implementing partners (HTSPE and Verde Azul):

a) The satellite mapping and inventory exercise
b) Capacity building of the local cadastral offices
c) Piloting a sound approach to area-wide registration of land rights

Activities under 'a' and 'b' are generic at the level of all selected 12 districts. However, activities under 'c' will be implemented only in selected priority areas in selected districts to address some hotspot issues related to expansion, requalification and regularization (Annex 1). The purpose of this area-specific interventions in priority rural areas is to pilot a sound approach to area-wide registration of land rights to individual rural households. Given the fact that activities 'a' and 'b' will be implemented everywhere in the 12 districts, the impacts to be identified at the beneficiary level in these priority areas will be the effects attributable to activity 'c' [Activity III].

5.2. Geographic coverage for the impact evaluation

A list of the 12 districts and the selection criteria they meet for Land Project activities is given in Table 2. Activity III is only implemented in selected priority areas in these districts. Since, collecting primary survey data from hotspots in all 12 districts was resource intensive and not practical, it was mutually decided by MCC/MCA and MSU to conduct the rigorous impact evaluation of the 'site-specific land intervention' only in two rural hotspot areas. Based on the scope of the activities planned and progress made by HTSPE in relation to the timeframe of the baseline survey, it was decided that the focus of the IE will be to evaluate the impacts of interventions targeted on hotspot issue of requalification / regularization in the following two districts—Mecufi in Cabo Delgado and Malema in Nampula. These two hotspot areas were selected for evaluation based on the following additional criteria which are critical for rigorous impact evaluation:

- Ability to identify comparison Aldeias to estimate the effects of the intervention in a rigorous and robust manner
- Indication that project interventions in hotspot areas will be implemented soon after the baseline survey and there will be enough time to observe outcomes and impacts before the end-line survey.

The geographic coverage includes 2 priority Aldeias in Malema and 3 priority Aldeias in Mecufi (Table 13). These were selected and prioritized by the district authorities (and HTSPE) based on some set criteria and were outside the control/influence of the impact evaluation team.

Table 13. Project intervention Aldeias for rural "hotspot" site specific activities under Pillar III

Mecufi (Cabo Delgado)	**Malema (Nampula)**
Maueia Muitua Ngoma	Cabo Miquitaculo Cabo Niquile

5.3. Research questions addressed by the IE

The goal of the intervention in the priority hotspot areas is to register or grant land use rights (i.e., land titles to long-term or perpetual-use rights) to individual households.[11] Initiatives to strengthen the property rights (e.g, issuance of DUATs) are generally designed to result in clearly defined rights that are enforceable, transferable, and of appropriate duration and scope. Economic theory holds that more secured tenure should lower land-transaction costs, lower the risk of expropriation or conflict and encourage more efficient land uses and land investment, and contribute to productivity improvement and land market development. More productive land should result in higher asset/land values and higher incomes for property owners. Over time, as land and financial markets develop formal land rights can also be used as collateral for loans.

But on the other hand, whether and to what degree these various impacts of more secure and transferable property rights based on economic theory are realized depends on local conditions such as market development, financial institutions, legal frameworks, and beneficiary income. The purpose of the rigorous IE design for the two rural hotspot areas is to precisely measure and monitor these impacts and assess the causality in effects outlined in the impact pathway. The key research questions to be addressed by our evaluation of Activity III in Malema and Mecufi are whether and to what extent the area-specific activities (as listed in column 1 of Table 12) leads to the various impacts as listed in column 3 of Table 14.

Table 14. Impact Pathway of Area-specific Activities in Rural Area.

Activities	Outcome	Impact indicators
• Digitized base maps for "priority areas" • Demarcated plots • Issuing DUAT for the plots	• Make the process simple, cost-effective and faster • Increased security of tenure	• Reduced incidents of conflicts • Increased new commercial enterprises and activities • Increased level of investments on land parcels • More active land markets • More effective/productive land uses • Increased off-farm opportunities (labor mobility) • Higher demand for DUATs

5.4. Identifying the comparison communities

[11] Initially, there were plans to conduct rigorous IE of the community land fund project (iTC) under Pillar III. However, based on the design of the iTC project and given the vast and diverse issues to be potentially covered by iTC, it was not feasible to do a rigorous impact evaluation of this component of Pillar III.

There are two things needed to implement the DiD IE design:
1. Identification of treatment and comparison sites, and
2. Data collection from both treatment and comparison sites before and after intervention.

The prioritized Aldeias listed in Table 16 are the potential pool of treatment sites for this IE. The units of impact observation will be households. Thus, households within the boundary of these listed Aldeias serve as the treatment group. The following strategy was used to identify sufficient number of comparison households to implement the DiD design.

The current strategy for Mecufi includes Maueia, Muitua and Ngoma as treatment Aldeias and Secura A, Secura B, Zaulane A, and Zaulane B as control Aldeaias. This is a deviation from our original design due to the change in the implementation plan. Originally, the implementation plan was to intervene the coastal side of all the seven Aledias, and leave the non-coastal side of each of the Aldeias untreated. Correspondingly, our original strategy was to select the same number of treatment households (from the coastal side) and control households (from the non-coastal side) in each of the seven Aldeias. After the baseline survey was completed, the intervention plan changed to complete treatment coverage of some villages (Maueia, Muitua and Ngoma) and leaving others (Secura A, Secura B, Zaulane A, and Zaulane B) as control due to the strong objection from its members on the original intervention plan. In light of the fact that the intervention plan changed after the baseline survey was completed, we don't have much choice but to salvage as best as we could. While the new plan is less ideal than the initial design, the four control villages are facing similar issues as the three treatment villages and will remain as control for the next few years. HTSPE's contract with the MCA ended in August 2013 (and indeed the whole MCA compact ended in September 2013). So naturally neither HTSPE nor MCA have plans to intervene in that area in a forseeable future. And it is also important to ensure that Cabo Delgado SPGC has no such plans before the completion of the end line survey.

In the case of Monapo, the impact evaluation strategy includes two treatment Aldeias and one control Aldeias. The treatment Aldeias (Cabo Miquitaculo and Cabo Niquile) were selected by the project team and we had no influence on that decision. There were very limited choices for a comparable control area in the Monapo district. The only close match was the community of Cabo Macassa. Therefore, we selected Cabo Macassa Aldeias as the control site. Like in the case of Cabo Delgado, it is important that Nampula has no plans to intervene in the control area before the completion of the end line survey for this evaluation project.

5.5. Sample size and sample selection

At the time of the planning of rural IE surveys, MCA had made a substantial reduction in the scope and coverage of the rural intervention. As of May 2011, HTSPE estimated to capture around 2000 agricultural parcels across all provinces, targeting to cover about 500 parcels in one district per province. The plan for Cabo Delgado and Nampula for rural LTR work is indicated in Table 13.

The sample size of the rural evaluation was dictated by the size of the targeted number of treatment parcels in Mecufi and Malema as conveyed to us at the time of planning this IE, and the logistics of doing the survey in limited time available before HTSPE planned their activities in the selected villages. In Mecufi district, Cabo Delgado, our initial understanding of the LTR timeline and scope was that work would start in the village of Muaria in August of 2011 and quickly proceed northward to the village of Muitua and cover only land on the coastal side of the road linking the district capital to the provincial capital of Pemba in the north (see Figure 2).

The plan was to cover all villages to be covered by the intervention along the coast, but given that the survey could only commence around August, and public announcements were already planned for early August in the village of Muaria, this southernmost village was excluded from survey coverage. To avoid any overlap between survey implementation and HTSPE intervention, the questionnaire was divided into 2 visits. Those sections that would have been sensitive to interventions were implemented first across all the villages.

Table 15. HTSPE plan for rural LTR work in Cabo Delgado and Nampula as of the time of the planning of the rural IE design

Province	*District*	*PA*	*Area*	*Size of area*	*Estimated number of parcels to be captured*
Cabo Delgado	Mecufi	Mecufi	Highway – Sea, northern part	25 km2	400-500
Nampula	Monapo	Monapo	Monapo Sede - Western part	100 km2	500

SOURCE: Interview with HTSPE staff, Tommy Kalms, **May 25, 2011**.

A listing was carried out in all villages and covered all households within them. Those households that owned parcels on the coastal side of the road were listed in the frame for treatment households; those with parcels only on the interior were listed in the frame for control households. The instruments used in the listing exercise are provided in Annex 1.

The number of treatment households selected in each village was calculated in direct proportion to the number of households listed in the frame for treatment households. An equal number of control households were then selected for each village from the frame of control households. The selection of households was done by systematic random sampling by the survey manager.

However, as noted in the previous section, the definition of control and treatment areas was changed post-survey. Due to a strong objection from community members, HTSPE proposed doing complete treatment coverage of a sub-set of villages (Maueia, Muitua and Ngoma) and left others as control villages (Secura A, Secura B, Zaulane A and Zaulane B). The ratio of sample size for the treatment and control villages in Mecufi that has been used for this study.

As with Mecufi district, the plan was to cover 400 treatment households and 400 control households. The survey was carried out under a very tight deadline because HTSPE was already scheduled to move in after a few weeks. Because the last 3 kilometers to Niquile could only be covered on foot and the residents widely dispersed, the enumerators had to have, at hand, precise instructions on the listing and sampling of households. It was decided that the 400 households be allocated between Cabo Miquitaculo and Cabo Niquile according to census information given by community leaders. This resulted in a 300-100 split between the two. Both communities had smaller administrative units called *celulas*. The distribution of the sample in Cabo Miquitaculo was done in proportion to the number of households listed within each *celula*.

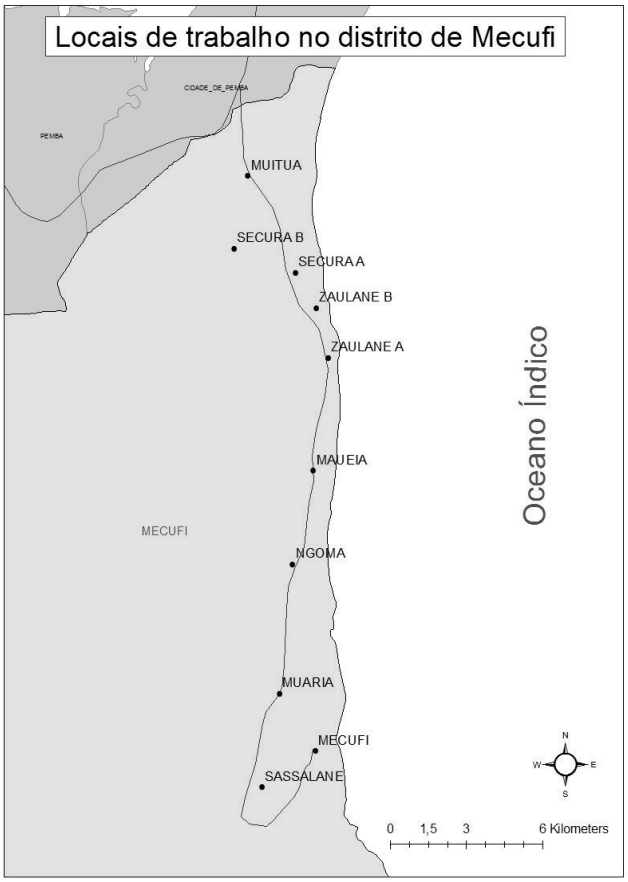

Figure 2 – Mecufi intervention area.

Source: Pinheiro, Andre. 2011.Actualização de Informação Cadastral na Zona Costeira de Mecufi. Report submitted by HTSPE/MCA-Mozambique/Verde Azul. February 2011.

In the case of Niquile, where the listing information could not be consolidated (owing to the highly dispersed population) before selection began, the 4 *celulas* were assigned 25 households each to facilitate the selection of households. The households were selected using systematic random sampling and the data were weighted using sampling weights.

In the Malema district, Nampula Province, the highest number of requests for DUATs came from those owning agricultural parcels in the low-lying areas along the Ligonha river. The river separates Malema district from Alto Molocue district in Zambezia. HTSPE indicated they would target this area. Most of the residents of this target area lived in the nearby communities of Cabo Miquitaculo and Cabo Niquile and so these were selected to be the treatment villages. There were very limited choices for a comparable control area. The only close match was the community of Cabo Macassa.

Table 16 summarizes the listing and selection information. The target number of interviews was not achieved in Cabo Delgado mainly due to absent respondents even after the second visit. In

Malema, the main constraint was that there were only 333 households in the final frame for control households.

Table 16: Number of households listed and selected by community and province.

Community	Total number of households listed	Total number of households with parcels in low-lying areas near the river	Total number of households interviewed	Number of households interviewed as % of households	Number of households interviewed as % of households with parcels in low-lying areas near the river
CABO DELGADO					
Ngoma	473		208	44.0	
Muitua	985		211	21.4	
Maueia	188		36	19.1	
Secura B	562		73	13.0	
Secura A	574		68	11.8	
Zaulane A	1,298		66	5.1	
Zaulane B	1,097		44	4.0	
Total	5,177		706	13.6	
NAMPULA					
Cabo Miquitaculo	718	465	297	41.4	63.9
Cabo Niquile	258	153	98	38.0	64.1
Cabo Macassa	473	333	316	66.8	94.9
Total	1,449	951	711	49.1	74.8

5.6. Power Calculation

As indicated in the previous sections, due to the extremely tight schedule between the time when the implementation plan was developed and the time when the plan was implemented and the challenge of small number of targeted parcels in both Malema and Mecufi, we had little choice with regard to the selection of treatment and control villages and number of households to be selected from each village. Moreover, our original sample design was further challenged by the fact that we had to switch between treatment and control areas due to the change of implementation plan. Given all the challenges, it is useful to conduct an ex post power calculation (shown below) using information from the baseline survey to assess whether we will still be able to accomplish a valid and rigorous evaluation of the rural hotspot project.

Formula (5) implies an important tradeoff between number of clusters and number of households per cluster sampled. For a given sample size, an increase in the number of households per cluster sampled increases the precision (i.e., reduction in MDE) much less

than an increase in the number of clusters sampled. Generally speaking, a relatively large number of clusters (e.g., 10 or more) is desired for an evaluation of a cluster-based intervention. For this reason, the small number of villages in the two program sites (7 in Mecufi district and 3 in Malema district) is a potential concern. One way to increase the number of clusters is to divide villages into sub-villages based on the assumption that households from different sub-villages have little interaction. It turns out that it is reasonable to divide the three villages in Malema district into 21 independent sub-villages. Specifically, the Cabo Miqitaculo Aldeias is divided into 11 subvillages (Chipaca A, Chipaca B, Murrapane, 25 de Junho, Metilili, 19 de Outubro, Nroposso, Mapecha, Lituli, 1 de Maio, and Pilani), the Cabo Niquile village into 4 subvillages (Namalelene, Nihoro, Mocuba, Chuhuro) and the Cabo Macassa village into 6 subvillages (Niessa, Euile, Murrosi, Murrunha, Uchequeche Namale). As a result, we have 21 clusters instead of 3 clusters to work with in the Malema district. The detailed distribution of sample by districts and by treatment status is listed in table 15.

Table 15. Sample distribution by communities and treatment status

Community	Sub-unit	Treatment	Control
Nampula			
Miquitaculo	Miquitaculo-Chipaca B	36	0
Miquitaculo	Miquitaculo-Murrapane	15	0
Miquitaculo	Miquitaculo-Chipaca A	32	0
Miquitaculo	Miquitaculo-25 de Jun	23	0
Miquitaculo	Miquitaculo-Metilili	58	0
Miquitaculo	Miquitaculo-19 de Out	12	0
Miquitaculo	Miquitaculo-Nroposso	32	0
Miquitaculo	Miquitaculo-Mapecha	14	0
Miquitaculo	Miquitaculo-Lituli	31	0
Miquitaculo	Miquitaculo-1 de Maio	25	0
Miquitaculo	Miquitaculo-Pilani	19	0
Niquile	Niquile-Namalelene	25	0
Niquile	Niquile-Nihoro	25	0
Niquile	Niquile-Mocuba	25	0
Niquile	Niquile-Chuhuro	23	0
Macassa	Macassa-Niessa	0	48
Macassa	Macassa-Euile	0	47
Macassa	Macassa-Murrosi	0	28
Macassa	Macassa-Murrunha	0	90
Macassa	Macassa-uchequeche	0	56
Macassa	Macassa-Namale	0	47
Total		395	316
Cabo Delgado			
Muitua		211	0
Maueia		36	0

Ngoma	208	0
Secura A	0	68
Secura B	0	73
Zaulane B	0	44
Zaulane A	0	66
Total	455	251

Following the traditional standard, we set the statistical power at 80% and the level of significance at 5%, which gives the multiplier value M_{j-2} to be 2.8 (corresponding to 2-sided hypothesis). Plugging M_{j-2}=2.8 and the other corresponding parameters for each respective district (see Table 16) into the MDE formula yields the standardized MDE (or MDE/σ) the respective district. Specifically, we will have (MDE/σ) = 0.28 for Malma and (MDE/σ) =0.35 for Mecufi, respectively. The smaller the MDE, the more powerful the design is. According to Duflo et al. (2006), a traditional norm is that a MDE of 0.3 is considered as "small", 0.5 as medium and 0.8 as big. Following this criteria, the design in both Nampula and Cabo Delgado is promising.

While the original sample design (small size and distribution) was not carefully designed due to the various reasons mentioned above, the MDEs for both Malma and Mecufi are reasonably small thanks to the small intracluster correlation coefficients. In other words, farmers from a given village (in the case of Cabo Delgado) or a given subvillage (in the case of Nampula) are fairly independent in physical asset endowment as well as economic variables.

Table 16. Power calculation by district

	Cabo Delgado	**Nampula**
M_{j-2}	2.80	2.80
P	3/7	15/21
J	7	21
N	101	34
ρ	0.017	0.014
MDE	0.35	0.28

Note: ρ is the mean intracluster correlation coefficients of a number of most relevant variables that were calculated based on the baseline survey data (Appendix table 2)

5.7. Baseline data collection

The baseline data were collected by interviewing the head of the households using a structured questionnaire. The questionnaire included more than 25 sections encompassing modules on:

- Household characteristics (demographic information by each member of the HH)
- Employment and sources of any other cash transfers
- Identification and list of all the parcels
- Land conflicts
- Rights to the land and perceptions of the risk
- Parcels rented out, rented in
- Characteristics of parcels
- Investments on land
- Perceptions about the DUAT, renting land and the land law
- Relative space occupied by crops in the plot

- Production and sales of basic food crops, cash crops, vegetables, fruits, nuts, etc., by season
- Input use by plot
- Agricultural practices
- Ownership of Assets
- Monthly expenditures
- Credit in the last 12 months
- Livestock and sub-products produced and sold in the last 12 months
- Consumption

Table 17. Number of households surveyed by type of community

Mecufi District, Cabo Delgado				Malema district, Nampula			
	HHs interviewed					HHs interviewed	
Aldeia	**Treatment**	**Control**		**Aldeia**	**Block**	**Treatment**	**Control**
Maueia	36			Cabo Miquitaculo	Chipaca A	32	
Muitua	211			Cabo Miquitaculo	Chipaca B	36	
Ngoma	208			Cabo Miquitaculo	Murrapane	15	
Secura A		68		Cabo Miquitaculo	25 de Junho	23	
Secura B		73		Cabo Miquitaculo	Metilili	58	
Zaulane A		66		Cabo Miquitaculo	19 de Outubro	12	
Zaulane B		44		Cabo Miquitaculo	Nroposso	32	
				Cabo Miquitaculo	Mapecha	14	
				Cabo Miquitaculo	Lituli	31	
				Cabo Miquitaculo	1 de Maio	25	
				Cabo Miquitaculo	Pilani	19	
				Cabo Niquile	Namalelene	25	
				Cabo Niquile	Nihoro	25	
				Cabo Niquile	Mocuba	25	
				Cabo Niquile	Chuhuro	23	
				Cabo Macassa	Niessa		48
				Cabo Macass	Euile		47
				Cabo Macass	Murrosi		28
				Cabo Macass	Murrunha		90
				Cabo Macass	Uchequeche		56
				Cabo Macass	Namale		47
Total	**455**	**251**				**395**	**316**
Overall	**1,417**						
Treatment	**850**						
Control	**567**						
Source: MCA/MINAG Rural Land Survey, 2011/12							

The survey had detailed sections for each of the outcomes to be evaluated, both intermediate and final outcomes. In addition, each of the survey households was geo-referenced for ease of locating them for the panel survey. In households that were male-headed with a spouse present, the spouse was the respondent for the livestock and food consumption modules. The survey was designed to take between 1 and 1 ½ hours to complete.

The baseline survey was implemented in September/October 2011 in Mecufi, Cabo Delgado and April/May 2012 in Malema, Nampula. If the head of the household was not present at the time of

the first visit, enumerators tried to make an appointment and returned again to interview the appropriate person within the time that the survey team was in the area. A total of 1,417 households were interviewed. The breakdown by province and treatment group area is shown in table 5.

5.8. Assessing the Validity of the Original Evaluation Design (As of May 2016)

5.8.1. Validity of the Evaluation Design for the Malema District

During the field trip in May, the MSU team also visited Malema (one of the two rural hotspot areas subject to rigorous evaluation) to check whether there has been any change in project implementation that undermined the original evaluation design. The key issue identified during the visit is that 5 out of the 15 treatment blocks did not receive any intervention by the end of the compact.[12] This finding has significant implication on the validity of the original evaluation design. In order to assess whether there will still be a rigorous evaluation using the original sample and the baseline survey data, we need to recalculate the MDE by accounting for the fact that five treatment blocks become control blocks.

Again, we use Eq. (5) to recalculate the MDE based on the updated information. The multiplier value M_{j-2} associated with the conventional power and level significance (80% and 5%) is 2.8. Plugging M_{j-2}=2.8 and the updated parameters (J=21, P=0.48, N=34, and ρ=0.014) into equation (5) yields the standardized MDE (or MDE/σ) =0.25. According to Duflo et al. (2006), a traditional norm is that a MDE of 0.3 is considered as "small", 0.5 as medium and 0.8 as big. The smaller the MDE, the better is the design. Following this criteria, we are fairly confident that the project in Malema can be rigorously evaluated if there is no further contamination in our sample. In other words, if the control units can remain valid control units before the endline survey is implemented, the baseline survey data and the data to be collected from the same households in 2017/2018 allow us to evaluate the impact of the land titling project in Malema.

5.8.2. Validity of the Evaluation Design for the Mecufi District

Compared to the Malema district, the situation in Mecufi district is much less clear. There has been a rumor that many parcels in the study areas in Mecufi were sold to investors and the original land owners were relocated to areas that are far away from their original villages. If this is indeed a case, it would be extremely challenging to collect data from those households who have moved. Unfortunately, this rumor has not yet been confirmed. During the field trip in May, 2016, we checked with land administrators in Nampula about this rumor, they were not able to confirm it at that time. On the other hand, based on our knowledge and impression from our early field trips to Mecufi, we wouldn't be surprised if many parcels were already sold. It is important that more information is gathered to decide whether a valid impact evaluation is possible in Mecufi.

[12] **These five sub-communities (highlighted in YELLOW in Table 15) are Mapecha, Nroposso, and Metilili in Cabo Miquitaculo communities, and Mocuba and Pilani sub-communities in Cabo Niguile communities.**

Table 15. Sample distribution by communities and treatment status

Aldeia	Blocks	Treatment	Control
Nampula			
Miquitaculo	Miquitaculo-Chipaca B	36	0
Miquitaculo	Miquitaculo-Murrapane	15	0
Miquitaculo	Miquitaculo-Chipaca A	32	0
Miquitaculo	Miquitaculo-25 de Jun	23	0
Miquitaculo	Miquitaculo-19 de Out	12	0
Miquitaculo	Miquitaculo-Lituli	31	0
Miquitaculo	Miquitaculo-1 de Maio	25	0
Niquile	Niquile-Namalelene	25	0
Niquile	Niquile-Nihoro	25	0
Niquile	Niquile-Chuhuro	23	0
Miquitaculo	Miquitaculo-Metilili	0	58
Miquitaculo	Miquitaculo-Nroposso	0	32
Miquitaculo	Miquitaculo-Mapecha	0	14
Miquitaculo	Miquitaculo-Pilani	0	19
Niquile	Niquile-Mocuba	0	25
Macassa	Macassa-Niessa	0	48
Macassa	Macassa-Euile	0	47
Macassa	Macassa-Murrosi	0	28
Macassa	Macassa-Murrunha	0	90
Macassa	Macassa-uchequeche	0	56
Macassa	Macassa-Namale	0	47
Total		395	316

Note: The rows highlighted by yellow color are the five treatment blocks that were not treated due to change in the implementation plan.

5.9. Implications and Next Steps on Evaluation of Activity III in Malema and Mecufi

A revised ex post power calculation shows that there will be a rigorous evaluation of Activity III in Malema district despite the fact several treatment communities were not treated. It is important to ensure that there will be no other land titling programs implemented in our study areas between now and the endline survey. We recommend the endline survey to be conducted in 2018/2019 to allow sufficient time for the real impact of the titling program to realize.

We are less enthusiastic about continuing the evaluation in Mecufi district. First, the ex post power calculation shows the value of MDE is much larger in Mecufi than in Malema (0.38 vs. 0.25) even with the assumption that the implementation exactly followed its original plan. Based on what happened in the Malema district as well as in the Monapo vila and the Nampula city, we are doubtful that the implementation did not deviate from its original plan. Second, there is a real possibility that many parcels have changed owners. If this turns out to be the case, it will be

extremely challenging to interview the original households if they have moved far away from their original locations.

If the budget is only sufficient to implement the endline survey for one of the two rural hotspot areas, we would recommend to choose the Malema district based on the power calculation. But if MCC is interested in evaluating Activity III in both districts, it is important to gather additional information from Mecufi to recalculate the statistical power and determine whether the investment in Mecufi is really worth it.

References

Deininger, K., Daniel Ayalew Ali, and Tekie Alemu. (2011). "Impacts of Land Certification on Tenure Security, Investment, and Land Market Participation: Evidence from Ethiopia," Land Economics 87(2): 312-334.

Di Teller, R., Sebastian G. and E. Schargrodsky (2007). "The Formation of Beliefs: Evidence from the Allocation of Land Titles to Squatters," Quarterly Journal of Economics, 122(1)209-41.

Duflo, Esthur, Rachel Glennerster, and Michael Kremer. (2007). "Using Randomization in Development Economics Research: A Toolkit" (with Rachel Glennerster, and Michael Kremer) in T. Paul Schults, and John Strauss (eds.) Handbook of Development Economics, Elsevier Science Ltd.: North Holland, 2007 Vol. 4, pp. 3895-62. (see also NBER Technical Working Paper No. 333, December 2006).

Purdon, Susan. 2002. "Estimating the impact of labor market programs", Department for Work and Pensions Working Paper No. 3, London, UK.

Field, Erica. (2007). "Entitled to Work: Urban Property Rights and the Labor Supply in Peru," *Quarterly Journal of Economics* 4(122): 1561-1602.

Ravallion, M. (2005). "Evaluating Anti-Poverty Programs," World Bank Policy Research Working Paper 3625.

Shackman, G. 2001. Sample size and design effect. Paper presented at Albany Chapter of American Statistical Association, March 24, 2001. http://www.albany.edu/~areilly/albany_asa/confweb01/abstract/Download/shackman.pdf

Annex 1: Definition of important terms and concepts in the context of the Land Project

Geographic areas: are basically "priority areas" that are facing some hot issues related to land that need urgent attention (e.g., conflict resolution, regularization, expansion, great demand that cannot be met with current capacity, etc.).

Bairros: Refer to a sub-set of a municipality with well-defined boundaries. They are similar to large neighborhoods (defined in terms of city blocks) in an urban area. These will be the unit of intervention for hotspot issues in the selected municipalities.

Villages: Refer to a sub-set of a district with well-defined boundaries in terms of inhabitants. These will be the unit of intervention for hotspot issues in the selected districts.

Hotspots: refer to (hot) issues that need to be resolved/addressed in a given geographic area. As such the geographic area to be identified for interventions may include 1-3 of the following hotspot issues.

- **Expansion:** This refers to the plan for expanding the area under a bairro based on a proper structural plan.

- **Requalification:** This is mainly a hotspot issue in urban areas that involves several steps with the end result being a restructured bairro that is properly zoned, roads are clearly marked, and each plot is demarcated and identifiable in the cadastral system with information on the name of the occupant(s), characteristics of the plot, demographics of the HH, etc.

- **Regularization:** Regularization (in the context of an urban setting) refers to the demarcation and delimitation of plots after an area is 'requalified.' Thus regularization is the follow-up step or the end result of requalification. Since the purpose of delimitation and demarcation is to register each plots in a cadastral system for potential DUATs, the municipality will not do this until they go through the 'requalification' process first.

Appendix Table 1: Intra-cluster correlation coefficients (ρ) of urban communities based on baseline data collected in 2011

	Monapo Vila		**Nampula**
Key Characteristics	Bairro level	Bairro level	EA level
Household size	0.0116	0.0159	0.0208
Gender of the head	0.0000	0.0041	0.0347
No. of parcels owned and currently in their possession	0.0192	0.0433	0.0654
Total household land size holdings (m2)	0.0000	0.0069	0.0016
Total No. of the parcels	0.0376	0.0978	0.1117
Total value of assets per hh (Mt)	0.0404	0.0025	0.0664
Participated in ag. Production	0.0075	0.0792	0.0619
Participated in livestock production	0.0000	0.0439	0.0451
Key outcome variables			
Total monthly hh non-food expenditure (Mt)	0.0000	0.0000	0.0000
Total monthly food consumption (Mt)	0.0004	0.0365	0.0524
Total monthly hh non-food expenditure using X items only	0.0000	0.0077	0.0000
Household Diversity Dietary Score (max.12)	0.0029	0.0303	0.0383
Total annual hh income including land rentals (MT)	0.0031	0.0018	0.0000
Participated in salaried income	0.0005	0.0009	0.0000
Participated in self-employment	0.0000	0.0365	0.0586
Participated in remittances	0.0126	0.0180	0.0424
Average of the key outcome variables	0.0060	0.0141	0.0260

Appendix Table 2: Intra-cluster correlation coefficients (ρ) of rural communities based on baseline data collected in 2011/2012

	Cabo Delgado	Nampula
Characteristics		
Total area of household landholding	0.006	0.007
Participation in maize production	0.001	0.008
Total value of assets per person	0.009	0.01
Total value of assets per adult equivalent	0.013	0.01
Sorg	0.001	0.005
Key indicators		
Total household income	0.006	0.003
Income per person	0.007	0.003
Income per adult equivalent	0.008	0.002
Total household expenditure	0.022	0.089
Total household expenditure p.c.	0.009	0.049
Total household expenditure per adult equivalent	0.017	0.057
Decision to make investment	0.025	0.043
Value of total investment	0.002	0
Hypothetical land value	0.004	0.0001
Hypothetical rental value	0.003	0.003
Perceived risk of losing land in the future	0.017	0.03
Willingness to pay for DUAT	0.05	0.016
Average	0.012	0.019