



# CONDUCTING RANDOMIZED CONTROLLED TRIALS (RCTs) TO EVALUATE THE IMPACT OF LAND AND RESOURCE GOVERNANCE SECTOR INTERVENTIONS: STRENGTHS, PRACTICAL CHALLENGES, AND BEST PRACTICE GUIDANCE

FEBRUARY 2022

# CONDUCTING RANDOMIZED CONTROLLED TRIALS (RCTs) TO EVALUATE THE IMPACT OF LAND AND RESOURCE GOVERNANCE SECTOR INTERVENTIONS: STRENGTHS, PRACTICAL CHALLENGES, AND BEST PRACTICE GUIDANCE

FEBRUARY 2022

USAID Contract Number: GS00F061GA

Communications, Evidence and Learning (CEL) Project

Authors:

Miguel Albornoz, NORC at the University of Chicago

Phoebe Bui, NORC at the University of Chicago

Lauren Persha, NORC at the University of Chicago

Suggested Citation: Persha, L., P. Bui, and M. Albornoz (2022) *Conducting Randomized Controlled Trials (RCTs) to Evaluate the Impact of Land and Resource Governance Sector Interventions: Strengths, Practical Challenges, and Best Practice Guidance*. Washington, DC: USAID Communications, Evidence and Learning (CEL) Project.

## DISCLAIMER

The author's views expressed in this publication do not necessarily reflect the views of the United States Agency for International Development or the United States Government.

## **CONTENTS**

<b>EXECUTIVE SUMMARY</b>	II
<b>I. INTRODUCTION AND BACKGROUND</b>	I
<b>II. SETTING THE STAGE</b>	5
Overview of LRG Sector Interventions	5
Learning Contributions from Prior Impact Evaluations	6
Knowledge Gaps	7
<b>III. THE WHAT, WHY, AND SO WHAT OF RCTs</b>	8
What are RCTs and Why are They Sometimes Viewed as Controversial?	8
RCT Contributions to Learning on Land and Resource Governance	10
Summary of Land Sector RCTs	10
Summary of RCTs by Broad Intervention Type and Outcomes Measured	17
<b>IV. PRACTICAL CHALLENGES, ETHICAL CONSIDERATIONS AND MITIGATING OPTIONS FOR RCTs OF LRG INTERVENTIONS</b>	20
Common Implementation Challenges and Mitigation Options	20
Ethics in Land and Resource Governance RCTs	26
RCT Design Considerations	28
Addressing Potential Risks and Harms	32
Cost-Effectiveness Analysis as an Add-on Consideration for LRG RCTs	34
Blended Methods	35
<b>V. BEST PRACTICE GUIDANCE</b>	36
When to Consider an RCT of an LRG Intervention	36
Best Practices for Implementing RCTs of LRG Interventions	37
<b>LITERATURE CITED</b>	40
<b>ANNEX I. INCREASE IN DEVELOPMENT RCTs OVER TIME</b>	44
<b>ANNEX II. RCT DESIGN CONSIDERATIONS AND DECISION-TREE</b>	45

# **EXECUTIVE SUMMARY**

## **BACKGROUND CONTEXT AND STUDY PURPOSE**

Land and resource governance (LRG) interventions often aim to make land access more equitable and to strengthen individual, household or communal rights to land and natural resources. LRG tenure strengthening and land formalization programming at USAID has focused on a range of interventions, including individual titling of private land and mapping and certification of individual or communal customary land rights. Such interventions span several broad categories, including those focused on land use planning and natural resource management, property rights and boundary clarifications, official rights recognition, land administration capacity building, awareness raising and sensitization on land rights and regulations, and legal, regulatory and policy dialogue, advocacy and reform.

Research and evaluations of LRG programs aim to fill existing knowledge gaps about LRG programming and theories of change, through carefully designed and rigorous studies. For many learning interests related to LRG interventions, impact evaluations (IEs) can provide one of the strongest ways to measure the impacts of the interventions with confidence and gain evidence-based learning on specific issues. Among the different types of evaluation and learning approaches, randomized controlled trials (RCTs) are often seen as an especially useful IE tool for evidence-based learning. However, RCTs to assess the impacts of LRG interventions, and of land sector programming in general, have been very uncommon.

This report seeks to help demystify RCTs for land sector programming, discuss some of the challenges and potential solutions for implementing LRG RCTs, and ultimately serve as a resource document that can help USAID to make informed decisions about whether, when, why and how to engage in supporting an RCT of a land sector intervention. The intended audiences for this report are USAID/Washington, Mission and other operating unit staff engaged in land sector programming.

## **RCTs IN DEVELOPMENT PROGRAMMING**

RCTs are a specific type of impact evaluation that use randomization to determine which beneficiaries will receive an intervention. Using randomization to control who will receive the intervention improves the rigor of the study and strengthens confidence in and the credibility of the findings from the evaluation. However, RCTs have often been viewed with skepticism within development sectors. A key point of discussion has been around potential ethical concerns, including perceived unfairness about who receives an intervention or how randomization is determined. Other concerns relate to the practicality of implementing RCTs of development programs, given coordination requirements with implementing partners and the potential need to harmonize the intervention and implementation approach with the RCT design. Despite these issues, many concerns about RCTs can be mitigated through good study design, or the addition of best-practice mixed-methods data collection.

## **THE RCT EVIDENCE BASE FOR LRG INTERVENTIONS**

Recent systematic reviews of evidence and learning from LRG programming have highlighted a need for additional credible evidence on the impacts of LRG interventions across a range of issues and sectoral interests. Much of the current knowledge base on LRG interventions has been obtained from study

designs that are typically considered less rigorous than RCTs. Some practitioners have explicitly called for additional RCTs in the sector to help fill important knowledge gaps.

To help characterize the existing RCT evidence based for LRG programs, we identified and drew on 14 published and unpublished studies of RCT evaluations of 10 land sector programs. These studies were conducted in 8 countries: Benin, Bolivia, Liberia, Mongolia, Mozambique, Tanzania, Uganda, and Zambia. We also identified an additional seven LRG sector RCTs that are registered as a trial in development or ongoing, but for which we found no published or grey literature reporting outcomes to date. We draw on these studies to take stock of current LRG RCTs, the implementation challenges they have experienced, and to examine ethical concerns around their use in the sector, with an eye towards summarizing lessons learned and best practices to mitigate such challenges and concerns.

Most of the current RCTs focus on land formalization programs in the form of titling or provisioning of certificates that codify and recognize customary use rights to land. Geographically, the RCT portfolio is heavily concentrated in sub-Saharan Africa. All of the 10 RCTs employed a cluster-randomized approach and randomize treatment at the level of blocks, villages, or herder groups, while most of them measure household-level effects. Eight used stratified (block) randomization and two used a phased or pipeline design. Two of the RCTs employed a more complex multi-arm treatment design to examine individual and joint effects of complementary interventions (for example, land titling and an intervention explicitly aimed to expand access to credit).

The more recent generation of LRG RCTs currently under development are also heavily focused on sub-Saharan Africa. Four of these seven use a cluster-randomized design, and two of them employ a multi-arm treatment design. We see more of a focus in this set on smaller-scale studies to examine the impacts of more targeted aspects of interventions, rather than a general look at the average effects of land formalization interventions on a wide range of outcomes across many different outcome categories. Thus, these newer “next generation” studies appear to be more narrowly focused on obtaining rigorous evidence about specific learning questions or causal mechanisms of interest, and they rely on innovative designs and randomization approaches to do so. For example, more attention is paid to less studied types of interventions, and there seems to be a greater focus on how land programming can be coupled with other interventions to achieve intended gender equity and poverty reduction goals.

More than half of these LRG RCTs measure the cumulative impacts of several different interventions rolled out as part of the same program (for example, sensitization on land rights, combined with mapping and rights recognition, and support to a local government land administration system). In terms of broad intervention types and impacts that these RCTs aim to measure, more than half investigated the impact of awareness-raising about land rights and changes to the legal framework as either the main or a secondary focus of a property rights intervention. Property rights and boundary clarification interventions, such as programs that conduct land mapping and facilitate land rights recognition through some form of documentation to the landholder, are also well represented. Some of the studies have some emphasis on land use and natural resource management. Few LRG RCTs to date have focused explicitly on impacts of land administration support or effects of legal or regulatory dialogue and advocacy. This makes sense intuitively, as these types of broader interventions may be less amenable to an RCT design.

In terms of the types of outcomes that existing LRG RCTs have focused on measuring, by far the most common is tenure security (9 out of 10) and land disputes/conflict resolution (7 out of 10). Other common outcome measures include land documentation uptake (4 out of 10), agricultural or livestock investment and productivity (4 or 3 out of 10, respectively). Shorter-term outcomes are most strongly represented. There is a notable lack of coverage on co-titling or specific measures of women's empowerment (such as women's bargaining power or intra-household decision-making). Other notable gaps include attention to the effects of LRG interventions on land rights knowledge, food security, land and labor markets and credit access. Land use or related natural resource management outcomes are sometimes addressed, depending on the nature of the intervention.

## **COMMON IMPLEMENTATION CHALLENGES AND MITIGATION OPTIONS**

Across the RCT studies of land sector interventions that we reviewed, four implementation challenges were most commonly mentioned by the researchers, as summarized in the table below. All of these situations are common potential challenges for impact evaluations that can affect confidence in and the validity of the impact results, although none of them are specific only to RCT designs.

### **COMMON IMPLEMENTATION CHALLENGES FOR LAND AND RESOURCE GOVERNANCE RCTs**

KEY IMPLEMENTATION CHALLENGES	EXAMPLES FROM PUBLISHED STUDIES	NUMBER (N = 10)	POTENTIAL MITIGATION STRATEGIES
<b>Timing of evaluation endline data collection</b> relative to when short and longer-term outcomes and impacts are expected	Timing between the intervention and the endline survey round may have been too short to capture the effects of participatory mapping on internal and external conflicts.  EXAMPLE: Reyes-Garcia (2012)	6	Plan from design phase to conduct multiple follow-up rounds of data collection.
<b>Delayed or incomplete program implementation</b>	Policy decisions by government led to a delay in the production of titles for treatment communities. (Ayalew Ali et al. 2014; 2016)  High transaction costs constrained the mobility of paralegals in treated communities (Mueller et al. 2015)  Local elites influenced community interest to participate in land documentation process (Knight 2013b; 2014)	6	Maintain close coordination with IPs throughout activity lifetime; Plan for flexibility on when the endline round of evaluation data collection takes place; Consider oversampling at baseline.
<b>Spillover effects (contamination)</b>  Intervention affects non-beneficiaries, including members of control group	Individuals in control villages were exposed to paralegals, due the close proximity between treatment and control villages.  EXAMPLE: Mueller et al. (2015)	3	Conduct logic exercise on the potential for spillovers and possible channels at design phase; structure the evaluation sample and data collection instruments to enable measurement of spillovers.
<b>Treatment non-compliance</b>  Units assigned to the treatment group do not receive treatment	Communities that had volunteered to participate ended up withdrawing after powerful elites engendered opposition to the program within their communities.  EXAMPLE: Knight et al. (2013b)	3	Plan for additional units in the treatment group from design phase; Maintain close coordination with IPs throughout activity lifetime; if appropriate, support M&E data collection by IPs during activity lifetime to understand reasons for non-compliance; structure follow-up data collection to enable additional learning on unintended effects.

Other practical challenges that were encountered included:

- Delays in receipt of government information needed to finalize the randomization process;
- Lower than anticipated take-up rates for the intervention;
- Unanticipated costs to beneficiaries to receive land documents that affected intervention roll-out and delivery of the documents;
- Challenges isolating the effects of particular intervention components of interest due to concurrent timing of different elements of an intervention package, or imbalance on key characteristics of treatment and control groups despite randomized assignment.

In addition, many RCTs fail to consider potential spatial bias in the sample design, which can result in treatment and control group assignment that is not truly randomized or balanced on underlying characteristics that can affect outcomes. Cluster-randomized RCT designs can be particularly vulnerable to this, as they often draw on geographic areas (such as villages), as the unit of randomization. While beyond the scope of this report, we note that studies increasingly call attention to a need for RCTs (and IEs of all designs) to make better use of spatial analysis methods to ensure that findings are not vulnerable to hidden biases due to spatial effects and/or the spatial location and distribution of units for the study.

### **ETHICS IN LAND AND RESOURCE GOVERNANCE RCTs**

Key ethical concerns for development RCTs often focus on issues of targeting and the vulnerability of potential beneficiaries. Within the land sector, where some interventions such as land titling or customary land formalization confer stronger legal rights to beneficiaries relative to non-beneficiaries, there may be concerns about the randomized selection of some groups of people to receive the intervention while others do not. However, mitigation strategies are often available to overcome these concerns. This can include using a phased or pipeline RCT design, in which the intervention plans to provide the same service(s) or intervention to the control group after the evaluation is conducted, as has been done for some land sector RCTs. In some situations, an RCT can actually be seen as the fairest approach to determine who will receive the intervention, such as in situations where program resources are limited and cannot be provided to all potential beneficiaries. When there are targeting concerns, a regression discontinuity design may also provide a viable alternative in some situations, where the treatment and comparison group for the evaluation are constructed around an eligibility cut-off to receive the program. Ultimately, decisions on fairness or related potential ethical concerns for a given RCT will depend on the type of intervention under consideration, the options available for how the randomized selection may be conducted, and the beneficiary and implementation context.

### **DESIGN CONSIDERATIONS**

Designing and implementing an RCT of a LRG intervention requires frequent communication, information-sharing and collaboration with implementing partners. To date, USAID has implemented two successful RCTs of land-sector interventions, in Zambia and Tanzania. This required close coordination among the evaluation team, implementing partners and local authorities, and careful planning and sensitization with IPs and local authorities as to the benefits of this approach from early on in the evaluation design phase. This level of coordination and sensitization also has budget implications for implementing partners and the evaluation team.

The timing of RCT design phase activities is important. Evaluation teams must be able to hold discussions with USAID and IPs while IPs are still in the process of designing their intervention and finalizing implementation details. This allows ample time for both teams to discuss if and how randomization might be feasible, the learning issues that USAID and IPs are interested in from an independent evaluation, and how intervention details or aspects of implementation can be planned to meet those objectives.

LRG RCTs must also determine the *options for how program benefits will be assigned to participants*. LRG interventions can be applied at several different levels, depending on the nature of the intervention. For example, services can be provided to individuals or households, or they can be rolled out at more aggregate levels, such as to neighborhoods, communities, municipalities, or larger administrative units. Many RCTs in the land sector are designed as ‘cluster-randomized’ RCTs, in which randomization into treatment or control groups is done at the level of a cluster, such as a village or neighborhood. The treatment itself is then provided to all eligible sub-units within the cluster, such as households or individuals. All of the 14 LRG RCT studies we reviewed employed this type of cluster-randomized RCT design, with the unit of assignment either as blocks of land parcels, villages, communities, or herder groups.

LRG RCTs must also determine the *method that will be used to randomize units into the intervention and control groups*. In a classic randomization design, all eligible units are randomly assigned to either treatment or control groups. This is often considered the most equitable method of randomization because all eligible units have an equal chance of receiving program benefits. If resources allow, an alternative approach is to use a pipeline, randomized roll-out or phase-in randomization design, in which all eligible units receive the program, but at different times. Another option is to provide either some or all program benefits to control groups after the end of the initial implementation period. For both of these alternatives, it is also important for evaluation teams and IPs to consider what time frame is feasible, ethical and practical (among potential other considerations) before treatment of the control group begins.

## **ADDRESSING POTENTIAL RISKS AND HARMS**

As part of ethical considerations, evaluation teams and donors should think through and discuss the possibility that program beneficiaries and/or non-participants may be harmed by the RCT, including the possibility for unintended negative consequences as a result of the RCT design. An example could stem from a situation where the RCT design results in many people being selected into the treatment group who would otherwise be uninterested in or unlikely to benefit from the intervention, such that take-up and/or measured impacts from the intervention are artificially low and dampen future policy interest in an otherwise beneficial program. In addition, what is the possibility the control group could be made worse off if prevented from accessing otherwise available services as a result of the RCT? The likelihood of both of these in practice can vary depending on the nature of the intervention and beneficiary context, and how randomization is conducted.

In addition to potential harms from the RCT design itself, evaluation team must also consider possibilities for the intervention under study to result in unanticipated negative effects for beneficiaries.

Evaluation teams should take these possibilities into consideration from the earliest stages of IE design in collaboration with IPs, and work with IPs to monitor potential harms to beneficiaries or control group study participants throughout program implementation. Here, it is possible that an RCT design may be better situated than other evaluation approaches to provide credible evidence on such unintended negative effects, if they exist, given the higher level of rigor and confidence in results that RCTs often can provide relative to quasi-experimental IEs or performance evaluation designs.

## **WHEN TO CONSIDER AN RCT OF AN LRG INTERVENTION**

Best practice ways to address concerns about RCTs include transparent communication with program implementers and related stakeholders about the reasons for randomization, potential benefits, and randomization options that might be available. It is also important for evaluation teams to provide sufficient information and seek buy-in about the potential for an RCT design from early in the evaluation design process. There is a need to maintain close involvement of these stakeholders throughout the planning. USAID can play an essential role by helping to facilitate this communication from early in the evaluation design phase, and by clarifying for all USAID's main learning priorities out of the evaluation.

### **RCTs of LRG interventions may be appropriate to consider when the following conditions are met:**

There is likely to be a demonstrable value-add to learning beyond what could be obtained from other impact evaluation alternatives, such as from quasi-experimental approaches. Will the knowledge gained through the additional rigor of the RCT be likely to advance current understanding of impacts from the intervention?

The intervention being considered for evaluation is well-defined and can be implemented with relatively low variability. In addition, IPs will not face overly burdensome challenges to maintain fairly standardized implementation of the intervention across different treatment units, and there are no inherent reasons to expect the intervention to vary considerably from place to place.

Randomized assignment of treatment is possible logistically and politically, and unlikely to cause overly burdensome challenges to program implementation.

Randomized assignment of treatment has a low probability of causing potential harms to beneficiaries or control group participants.

Randomized allocation of the intervention is not likely to reduce the treatment effects on participants due to variation in their ability or interest to benefit from the treatment.

The planning for the intervention itself is still at an early enough stage, such that an evaluation team has sufficient time to scope out potential RCT approaches, discuss options collaboratively with IPs and USAID, finalize the design and any schedule or implementation requirements for the intervention and collect the evaluation baseline data before the planned start of the intervention. In practice, this typically means that the evaluation scoping work should start prior to or concurrent with the program award to IPs.

Other questions that evaluation teams and donors should ask while considering RCT options include:

- *What implementation and other assumptions are associated with an RCT approach for a given intervention and evaluation design context?*
- *What is the specific treatment(s) or intervention(s) that could be tested through an RCT evaluation? (Particularly important to clarify when the program under evaluation has multiple components)*
- *What are the possibilities for conducting a targeted RCT of one or more components within a broader program, or focusing on other elements of the intervention or how it is delivered that could help fill specific knowledge gaps about intervention effectiveness, causal mechanisms or impacts for different sub-populations?*
- *What is the potential added-value of the RCT from an evidence-based learning perspective, within the context of the current relevant knowledge and evidence base?*
- *How do the potential learning contributions mesh with USAID's learning priorities for the evaluation, and the potential contributions to the broader evidence base about the intervention or its theory of change?*

## BEST PRACTICE GUIDANCE

### FOR USAID AND RELATED DEVELOPMENT AGENCIES

- Develop clear learning priorities and articulate evidence needs for an independent evaluation, either with an evaluation team or prior to engaging the team. This can include visioning on how rigorous evidence-based learning from an evaluation might help to inform future programming decisions;
- Facilitate frequent and open communication and information-sharing between evaluation teams and IPs for the intervention under evaluation, and set budget expectations for this information-sharing and engagement across both teams from early on;
- Begin facilitating evaluation scoping and design work early in intervention planning.

### FOR EVALUATION TEAMS

- Obtain a detailed understanding of the implementation context and intended program design from earlier in intervention planning;
- Conduct frequent outreach and clear communication with donors and IPs about potential design options and seek buy-in from early stages;
- Work to achieve a common vision and set of goals for the RCT across the various partners, incorporate learning interests from IPs where possible, and help to build a common understanding of the evaluation design and objectives;
- Aim to avoid requiring IPs to operate in contexts that are less familiar to them, or to implement the intended program to those who typically would not be prioritized for the planned intervention;
- To the extent possible, ensure that RCT designs employ methods that also help to understand why programs work or not and pathways to impact, rather than just obtaining a statistical measure of the intervention's effects. Often this entails integrating qualitative and non-experimental data collection and research methods into the study, and ensuring that these components draw on diverse expertise across disciplines;
- Employ RCT approaches and sample designs that anticipate the possibility for spillover effects, partial treatment, or other threats to validity. Think through potential channels for spillovers from the intervention. If logic suggests the spillover potential will be high, aim for a design that enables testing for this;
- Consider ways to explicitly address heterogeneous treatment implementation, and ensure a design is sufficiently powered to examine heterogeneous treatment effects on different sub-populations of interest among the beneficiary population (for LRG interventions, this could include, for example, female-headed households; women within households; poorer or otherwise more vulnerable households; other traditionally marginalized groups);
- Anticipate that obtaining buy-in and adherence by IPs and government or other stakeholders will likely require an iterative process, much advance planning, and working collaboratively with all partners to explain the process, weigh trade-offs and ensure feasibility;
- Recognize that RCT approaches do often impose some additional constraints on IPs with respect to how they implement a given intervention, to maintain the integrity of the RCT. Work with IPs to discuss potential implementation permutations, and explore possibilities to work with this;
- When possible, use active time on the ground with the IP team in-country, for example through scoping visits, data collection preparations and data collection itself to firm up relationships and continue to strengthen collaboration throughout the evaluation;
- Maintain proactive and frequent communication with IPs regarding implementation progress once the intervention is underway;
- Involve IPs in all stages of the design and implementation of the RCT but seek to minimize the additional burden on IPs to the extent possible.

### FOR IMPLEMENTING PARTNERS

- Seek to understand the learning benefits of an RCT evaluation of the intervention and what the approach will entail;
- Consider how a rigorous evidence-based approach to learning from evaluation activities could be leveraged to provide additional learning for implementation or to help make course corrections as implementation proceeds;
- Work collaboratively with the evaluation team to communicate, raise any concerns, share information, and help ensure coordination between the program implementation and the RCT design and implementation;
- Use knowledge of the RCT to help work with and keep other program stakeholders, such as local government authorities, informed about the study;
- Aim to follow the agreed RCT protocol with respect to program implementation;
- Be proactive about reaching out to the evaluation team to discuss options when potential complications, delays, or other changes to planned implementation arise, or otherwise unexpected issues occur over the course of implementation that might be relevant for the RCT.

# I. INTRODUCTION AND BACKGROUND

## BACKGROUND CONTEXT AND STUDY PURPOSE

This document has been commissioned by the United States Agency for International Development's (USAID) Land and Resource Governance Division in the Bureau for Democracy, Development and Innovation (USAID/DDI/LRG). It responds to a request by USAID to undertake secondary research to examine strengths and weaknesses of designing and implementing randomized controlled trial (RCT) approaches to evaluate the impact of land and resource governance (LRG) interventions, including drawing on the current body of land sector RCTs to take stock of practical challenges in implementing RCTs, typical ethical concerns and best practices to proactively mitigate them.

In doing so, the report aims to serve as a resource to help USAID Missions and Washington make informed decisions about: (1) whether, when and how to engage in supporting an RCT of a LRG intervention (2) the potential value-add of doing so; and (3) the steps USAID staff can take to help facilitate the design and implementation of RCT impact evaluations of land sector programming. The primary intended audiences for this report are USAID/Washington, Mission and other operating unit staff engaged in land sector programming. The report concludes with recommendations for USAID, impact evaluation teams, and implementing partners (IPs) of USAID interventions.

## WHAT ARE LAND AND RESOURCE GOVERNANCE INTERVENTIONS?

USAID conceptualizes LRG as “the bundle of rules, rights, policies, processes, institutions and structures created to manage the use, allocation of, access to, control, ownership, management, and transfer of land and land-related natural resources” (Stevens et. al. 2020).<sup>1</sup> Land tenure, or “the relationship that individuals and groups hold with respect to land and land-based resources, such as trees, minerals, pastures, and water”, is an important element of LRG. Land tenure systems “define the ways in which property rights to land are allocated, transferred, used, or managed in a particular society” (USAID).

LRG interventions typically aim to make land access more equitable and to strengthen individual, household or communal rights to land and natural resources. Building on long-standing economic theory regarding the role of property rights in improving household livelihoods, especially in rural agricultural settings, LRG interventions in the development space work from the assumption that when individuals or households have stronger land rights and tenure security over land they use, they will have a greater incentive to invest in the land in ways that enhance its productivity, value and their own land-based incomes. Over the longer term, households' land investments and improvements to their agricultural productivity are also anticipated to improve their food security and overall economic wellbeing.

---

<sup>1</sup> Stevens, C., Panfil, Y., Linkow, B., Hagopian, A., Mellon, C., Heidenrich, T., Kulkarni, N., Bouvier, I., Brooks, S., Lowery, S., & Green, J. (2020), Land and Development: A Learning Agenda for Land and Resource Governance at USAID. More broadly, governance can be thought of as “the institutions, structures, and processes that determine who makes decisions, how and for whom decisions are made, whether, how and what actions are taken and by whom and to what effect” (Graham et al. 2003, Lockwood et al. 2010, Bennett and Sattersfield 2018). Institutions are the formal (laws, policies, tenure systems) and informal (social norms, prevailing power structures) rules that shape human interactions and that guide or constrain actions (North 1990).

LRG tenure strengthening and land formalization programming at USAID has focused on a range of interventions, including individual titling of private land and mapping and certification of individual or communal customary land rights. These programs often include multiple components in addition to documenting, mapping, and registering land rights and ownership, such as awareness building on land rights, supporting land use planning efforts in communities, strengthening local level land governance, administration and institutions, and addressing related issues of women's empowerment and adverse gender norms around land rights and access.

Much prior work has summarized the underlying theory, assumptions and causal pathways that illustrate how LRG interventions may lead to desired development outcomes and impacts over the short and long term.<sup>2</sup> Over the short and medium term, these outcomes range from improved land tenure security, possession of formal documentation of land rights and fewer land conflicts, to increased land investment and agricultural productivity and more sustainable land use. Several other outcomes are also anticipated in certain contexts, including improvements in food security, intra-household decision-making, women's empowerment, credit access, and land rental markets. Over the longer term, evaluations of LRG interventions seek to understand the extent to which LRG interventions lead to key development impacts, including poverty reduction and economic growth.

Many existing resources provide an understanding of how LRG interventions fit into the broader development space and the body of evidence to date has provided growing support for some key elements of LRG causal chains (for example, positive effects on farm-based investments and on women's empowerment). The current body of evidence also highlights some key knowledge gaps in the LRG sector, such as a weak or mixed evidence base for a link between improve tenure security via tenure strengthening interventions and positive effects on agricultural productivity or income. To date, rigorous evidence from longer terms studies on higher order impacts, such as poverty reduction or broader land use change and environmental effects, is also lacking.

Research and evaluations of LRG programs often aim to fill existing knowledge gaps through carefully designed and rigorous studies across a range of contexts. In addition, studies aim to better grapple with the often complex and highly nuanced causal pathways and interactions that influence outcomes of LRG interventions. For example, rigorous land sector evaluations increasingly seek to better understand how beneficiary and context factors, variations in the intervention itself and other features interact to shape different outcomes and impacts in different settings. Evaluations also increasingly aim to characterize how impacts may differ for different sub-populations of interest within a given development program (such as female household members, poorest households, or other particularly vulnerable or marginalized groups within communities), and the reasons why.

## WHAT ARE RANDOMIZED CONTROLLED TRIALS?

For many learning interests related to LRG interventions, impact evaluations (IEs)<sup>3</sup> can provide one of the strongest ways to measure the impacts of the interventions with confidence and gain evidence-based

<sup>2</sup> For example, see Meinzen-Dick et. al. 1997; Deininger 2003; Place 2009; Meinzen-Dick 2009; Holden and Ghebru 2016; Lawry et. al. 2017; Higgins et. al. 2018; Lisher 2019; Meinzen-Dick, Quisumbing, Doss et. al. 2019.

<sup>3</sup> USAID's Evaluation Policy defines impact evaluations as those that "measure change in a development outcome that is attributable to a defined intervention. They are based on models of cause and effect and require a credible and rigorously defined counterfactual to control for factors other than the intervention that might account for the observed change." (USAID 2016).

learning on specific issues. IEs make use of counterfactual logic (in other words, *what would have happened in the intervention communities if the intervention had not been implemented?*), and the designs of these evaluations enable researchers to attribute observed changes in outcomes to particular development interventions with greater confidence (White 2010; USAID 2011). The number of IEs in the development space has grown exponentially over the past two decades as IEs have risen to the forefront of methods that can provide credible evidence on the effectiveness of development interventions. IEs have provided decision-makers with important evidence-based learning that has been used to inform their development policy and programming decisions on what works, for whom, under what context conditions, and why (Sabet and Brown 2018).

Among the different types of evaluation and learning approaches, RCTs are often seen as an especially useful tool for evidence-based learning because of their stronger ability to isolate and measure effects of development programs that can confidently be attributed to the program itself and not to confounding factors. Relative to quasi-experiments and non-experimental approaches such as pre-post performance evaluations, RCTs can in theory provide the most unbiased estimate of causal impacts that truly result from a given program or intervention of interest (Ravallion 2020). For these reasons, they are often held up as the “gold standard” for evidence-based development learning (Banerjee, Duflo and Kremer 2016). However, RCTs to assess the impacts of LRG interventions, and of land sector programming in general, have been exceedingly uncommon to date. For example, a recent systematic review of the effects of increased land tenure security identified 59 rigorous studies to draw from, but just two of those were RCTs.<sup>4</sup> This report aims to help demystify RCTs for land sector programming, discusses some of the challenges and potential solutions for implementing RCT evaluations of land sector programming, and ultimately serve as a resource document that can help USAID to make informed decisions about whether, when, why and how to engage in supporting an RCT of a land sector intervention.

## **REPORT ROAD MAP AND LITERATURE SEARCH STRATEGY**

Below, we summarize the search strategy used to identify relevant published and unpublished RCT studies of LRG interventions. Section II provides a brief overview of LRG interventions, priority learning interests, learning contributions and key knowledge gaps. This includes a review of key strengths and weaknesses of RCTs for development learning from the perspective of LRG-specific learning needs. Section III summarizes the role of RCTs in development learning, perceived controversies, and the contributions of existing RCTs to learning on LRG development effectiveness.

Section IV discusses practical challenges and ethical considerations in designing and implementing RCTs of LRG interventions and introduces options to mitigate such concerns. Specifically, we ask: *what are particular challenges to conducting RCTs in the LRG sector that may stymie their wider implementation? To what extent are ethical concerns a limiting factor, and what lessons can be learned from existing land sector RCTs studies on how to potentially avoid or mitigate such concerns?* Section V provides best practice recommendations for USAID, evaluation teams and implementing partners.

---

<sup>4</sup> Higgins, D., Balint, T., Liversage, H., and P. Winters. 2018. Investigating the impacts of increased rural land tenure security: a systematic review of the evidence. *Journal of Rural Studies*. 61:34–62.

## LITERATURE SEARCH STRATEGY

We used a broad search strategy to identify relevant published and unpublished literature on LRG RCTs for this report. Prior land sector reviews have found few published RCTs of LRG interventions to date (for example, see Higgins et. al. 2018; Snistveit et. al. 2016). Some RCTs of LRG interventions are still ongoing, published only as working papers or donor-funded reports, or otherwise have not yet made their way into the published literature. As a result, we aimed to cover a range of potentially relevant sources beyond academic journal article databases. Our strategy included: 1) a systematic search of the Scopus database of academic literature; 2) a systematic search of relevant evidence hubs<sup>5</sup>; and 3) a targeted search of publication collections available through key practitioner and research organization websites (Global Land Alliance, Gates Open Research, Land Portal, World Bank DIME).<sup>6</sup>

We chose broad subject and evaluation key terms to cast a wide net. Our initial search on 3ie's evidence hub yielded 25 documents.<sup>7</sup> Our initial search on Scopus yielded 2,386 articles ("land" and "randomized"; "property rights" and "randomized"), and was further narrowed through the use of additional terms ("tenure"; "rights"). The searches were conducted in February 2020, and resulted in 484 documents for abstract screening. We did not restrict articles to a particular time period, although in practice we did not find any published LRG RCTs prior to 2010.

After accounting for duplicate references across databases, 14 studies out of those screened met our criteria for inclusion: studies must (1) evaluate a land and resource governance sector intervention, (2) use an RCT approach for causal identification and to estimate impacts of the LRG intervention, and (3) provide sufficient details on study design and outcomes measured, either through published papers, reports, or through an evaluation registry. Together, these 14 studies reported results from RCT evaluations of 10 land sector programs conducted in 8 countries: Benin, Bolivia, Liberia, Mongolia, Mozambique, Tanzania, Uganda, and Zambia. Twelve of the studies reported endline impacts of the intervention, and two reported interim findings prior to endline data collection. Four of them reported on effects for an intervention that had been cancelled or could not move forward as planned due to implementation issues. As a result, the focus for those studies was on secondary issues of interest.

In addition to these 14 studies of 10 ongoing or recently completed interventions, our search of the AEA RCT registry identified seven additional LRG sector RCTs that are currently registered as in development or ongoing, but for which we found no published or grey literature reporting outcomes to date.<sup>8</sup> These appear to be RCTs in progress for an additional seven LRG interventions in Bangladesh, Burkina Faso, Democratic Republic of Congo, Ghana, Philippines, Rwanda, and Uganda. We list these studies in Tables 3 and 4 below (see Section III).

Our search also yielded several studies that are broadly applicable for contextualizing LRG sector RCTs, including relevant systematic reviews, quasi-experimental evaluations, and evaluations of land, agriculture

<sup>5</sup> These include: 3ie's development evidence portal ([3ieimpact.org](https://3ieimpact.org)), MCC's evaluation repository ([mcc.gov/our-impact/independent-evaluations](https://mcc.gov/our-impact/independent-evaluations)), USAID's development experience clearinghouse ([dec.usaid.gov](https://dec.usaid.gov)), Campbell Collaboration systematic reviews ([campbellcollaboration.org](https://campbellcollaboration.org)), and the AEA RCT registry ([socialscienceregistry.org](https://socialscienceregistry.org)).

<sup>6</sup> We also conducted a targeted search of published papers by researchers and development economists who have conducted LRG impact evaluations in the past. Our search appears to have identified more land sector RCTs than recent similar efforts, but as many of these studies are in the donor sphere, where results may remain primarily as unpublished grey literature or donor-funded reports, we likely have not captured all potentially relevant work.

<sup>7</sup> Search string combinations: "land" or "tenure" or "rights" and "RCT" as the evaluation method.

<sup>8</sup> Of these, we excluded one study from additional discussion due to insufficient information in the AEA registry to characterize the study, such as information on randomization approach, primary outcomes of interest and country of implementation.

and forestry or other natural resource programs. We drew on some of this broader literature to inform our discussion in remaining sections of this report.

## II. SETTING THE STAGE

### OVERVIEW OF LRG SECTOR INTERVENTIONS

LRG programming encompasses several different types of interventions which may be implemented individually or more commonly as part of a broader package of programming. Theories of change for LRG interventions are necessarily program-specific, while programming logic for the sector has iterated over time as new learning is obtained.<sup>9</sup> The overarching logic frameworks for anticipated results are generally grounded in long-standing development theories around the linkages between particular types of LRG interventions, beneficiary behavior change, and related pathways to outcomes and impacts. A high-level illustration of this comes from USAID's 2020 "roadmap" of LRG interventions, which baskets LRG programming into five categories of interventions and highlights how these interventions are broadly conceptualized to relate to key shorter-term outcomes and high order development impacts over time (Figure 1).<sup>10</sup>

*Figure 1. USAID Roadmap for Land and Resource Governance Interventions*



(Source: Stevens et. al. 2020, adapted from GLTN and IFAD Theory of Change)

<sup>9</sup> For some recent examples, see Lawry 2017 and Higgins 2018.

<sup>10</sup> Per USAID, the roadmap is "intended as a general guide on the logical and temporal relationship between LRG programs and ultimate development objectives to end the need for donor assistance" (Stevens et. al. 2020).

## **LEARNING CONTRIBUTIONS FROM PRIOR IMPACT EVALUATIONS**

There is currently a large evidence base on the effects of a range of LRG interventions, which spans decades of research. Much of that knowledge base was derived from quasi-experimental, pre-post studies, qualitative research and other approaches that are typically considered to provide less rigorous evidence than RCTs, because these other study designs cannot measure causal impacts with confidence in the same way that RCTs can. Several recent systematic reviews of evidence and learning from LRG programming have highlighted a need for additional credible evidence on the impacts of LRG interventions across a range of outcomes and sectoral interests (for example see: Higgins et al. 2018; Lawry et al. 2017; Meinzen-Dick et al. 2019).

Investments in building the LRG knowledge base have come from many different sources. USAID alone has produced a substantial body of research and evaluation work on LRG programs, with some 150 research products produced since 2003. This includes eight impact evaluations (two of which are RCTs) and eight performance evaluations (PEs) of LRG interventions (Stevens et. al. 2020). MCC's land and property rights portfolio has similarly produced at least five IEs and five PEs in recent years.<sup>11</sup> The World Bank's Gender Innovation Lab has recently completed or is currently conducting RCTs of at least five LRG interventions in Benin, Ghana, Tanzania, Rwanda and Uganda, while other land-related studies may be underway through other research units.<sup>12</sup>

The growth in rigorous impact evaluation work of LRG interventions over the past decade or so is encouraging, and mirrors a growth in IEs across development sectors more broadly over the same time frame (Sabet and Brown 2018). However, RCTs of land sector interventions are still very uncommon. For example, in Higgins et. al.'s (2018) recent systematic review, the authors identified 59 robust studies examining land tenure security effects, primarily via land formalization programs or similar interventions. However, only two of the 59 studies they identified were RCTs.<sup>13</sup>

Our search results for this report also suggest there are a number of RCTs in the land sector that appear to be in progress but had not published results by the time of report writing. Thus, the body of studies in progress is likely greater than the small set of LRG RCTs with impact results that is currently available in the published literature. Our search, inclusive of donor-funded evaluation reports and other unpublished work, identified 14 RCT studies of 10 ongoing or recently completed LRG interventions, and another set of RCTs in progress for an additional seven LRG interventions (results not yet available). See related tables and additional discussion in Section III below.

Learning from the current body of LRG evaluations has been instrumental for characterizing patterns of outcomes and shorter-term impacts across a range of development issues;<sup>14</sup> understanding how different sub-populations of interest, such as women, poorer households or members of different ethnic

---

<sup>11</sup> See: <https://www.mcc.gov/resources>.

<sup>12</sup> Given the strong emphasis on RCT designs for World Bank program impact evaluations, we do not assume this is an exhaustive list.

<sup>13</sup> In their review of land tenure interventions, primarily focused on land titling and formalization efforts, Tseng et al. 2021 similarly find RCTs to be uncommon in the sector.

<sup>14</sup> These include: tenure security; agricultural productivity; land investment; credit access; land conflict; land markets; income and other aspects of household economic wellbeing; food security and/or nutrition; women's empowerment; and environmental impacts (see Higgins et al. 2018).

groups, may be differently affected by certain interventions such as formalization of customary land rights; and refining theories of change.

## **KNOWLEDGE GAPS FOR LRG INTERVENTIONS**

The current body of evidence has also helped to bring into focus key knowledge gaps. For example, existing studies and systematic reviews have not provided consistent evidence on whether and under what conditions land tenure formalization and other forms of tenure security strengthening can indeed lead to core outcomes in LRG theories of change (such as increased land investments), and the reasons why or why not. There are also gaps in knowledge around the extent to which impacts might vary for different populations of interest or in different context conditions (Fenske 2011; Lawry et. al. 2017; Higgins et. al 2018). Key knowledge gaps include:

- Whether, to what extent, and for whom does land tenure formalization or other forms of tenure security strengthening lead to increased land investments, agricultural productivity, access to credit, and household economic gains? – each of which constitute core elements of LRG theories of change but to date have not been consistently evidenced in several existing studies or in systematic reviews of the sector<sup>15</sup>
- What are the mechanisms by which land sector interventions achieve their results?
- How do results from different land sector interventions vary for different populations or based on geography, and how can this be used to improve targeting and cost-effectiveness of land sector development programming?
- How much time might need to accrue for impacts to be realized at scale?
- In what contexts might tenure formalization interventions *not* be the most efficacious solution, and why?

Recent summaries have also called for LRG IEs to do more to: (1) expand their geographic focus beyond contexts in sub-Saharan Africa, (2) build the evidence base for the effectiveness of other types of land formalization interventions beyond those that provide individual or household-level certification or titling, such as communal land titling efforts, (3) examine other baskets of LRG interventions beyond land formalization programs, such as interventions that support land use planning, aim to improve the effectiveness of land administration systems or provide land rights sensitization and legal aid programs, and (4) understand how and when land sector interventions should be coupled with those in other sectors, such as formalized credit or market systems support, to improve their ability to unlock broader development objectives (Lawry et. al. 2014; Higgins et. al. 2018). USAID and other donors recognize the need for additional rigorous studies to fill these and related key knowledge gaps<sup>16</sup> on the extent to which LRG interventions are likely to lead to their intended impacts, for whom, and under what conditions (Stevens et. al. 2020).

The vast majority of rigorous quantitative evidence of LRG interventions to date has relied on quasi-experimental approaches to identify intervention impacts, rather than RCTs. Some practitioners have explicitly called for additional RCTs to fill important gaps (see Higgins et. al. 2018). RCTs are particularly

---

<sup>15</sup> Scholars have also long pointed to inconsistent tenure security and outcome definitions, measurements and approaches across studies as a potentially key reason for the still inconclusive body of evidence with respect to tenure security-agricultural productivity linkages (Place, 2009; Arnot et al., 2011; Fenske 2011; Ghebru and Lambrecht 2017).

<sup>16</sup> Including a lack of attention to certain key outcomes from land sector programming, such as food security, and insufficient attention to intra-household differences in outcomes (Stevens et al. 2020).

well-suited to provide rigorous evidence to help fill several of the currently recognized knowledge gaps, including on:

- Longer-term impacts of LRG interventions;
- Causal mechanisms and the interactions of interventions, mechanisms and context to produce outcomes; and
- Attention to heterogeneity of treatment effects (for example, differences in impacts for men and women, wealthy and poor, and so on).

### III. THE WHAT, WHY, AND SO WHAT OF RCTs

#### WHAT ARE RCTs AND WHY ARE THEY SOMETIMES VIEWED AS CONTROVERSIAL?

Impact evaluations make use of a counterfactual that is obtained through a comparison or control group. The comparison or control group does not receive the intervention, but otherwise must be similar to the group receiving the intervention (the treatment group) (White 2013). RCTs are a specific type of impact evaluation that use randomization to determine which beneficiaries will receive an intervention. To do so, evaluation teams typically work with IPs to create a pool of units (for example, individuals, households, or villages, depending on the nature of the intervention) that are eligible to be a part of the planned intervention. Then, the team randomly selects which units will actually receive the intervention or not, out of that larger pool. The group that was randomly selected to receive the intervention is referred to as the treatment group or the intervention group. The units that are randomly assigned not to receive the intervention serve instead as the control group for the RCT impact evaluation. This group serves as the counterfactual. A simplified way to think of this is that because the units in this control group were similar to the group that was selected to receive the intervention, and the only reason they are not part of the intervention is due to the random selection process, this group provides credible information on what would have happened to the intervention group if they had not received the intervention. Using randomization to determine who will receive the intervention removes the likelihood of systematic

#### WHAT IS SELECTION BIAS?

*Selection bias can threaten the validity impact evaluation results. It is present when the group that receives an intervention is systematically different from the comparison group in ways that may also affect their outcomes from the intervention. For example, let's say an intervention selects only the least productive farmers to receive a farmer training program. Comparing evaluation results from those farmers to other farmers that did not receive the program may give a biased understanding of how the farmer training program affects yields, because other characteristics of low producing farmers or their farms could also affect the success of the farmer training program.*

*If not properly accounted for in an evaluation design, selection bias can make comparisons between the group that receives an intervention and a comparison group less valid. It can also reduce confidence in the estimates of program impacts obtained by an evaluation, as those impacts may not be solely due to the intervention itself. It is often easier to avoid selection bias through RCTs than with other evaluation designs.*

differences between the treatment and control groups, thereby overcoming threats of selection bias.

RCTs are viewed as a superior evaluation approach because of their ability to overcome selection bias and generate more precise and less-biased estimates of the effects of interventions. Researchers also have greater confidence that any differences in program outcomes or higher-level impacts between the two groups are indeed attributable to the intervention itself, and not to confounding factors (White 2013). Conceptually, RCTs can also be easier to communicate to a non-technical audience or to policy-makers who may make programming decisions based on evaluation results (White 2013). Analysis of RCT datasets is also typically more straightforward compared with quasi-experimental designs, which generally must use more complicated statistical approaches to render treatment and comparison groups similar to each other and to check the credibility and robustness of the results.

Commonly cited criticisms of RCTs include their potentially limited external or internal validity<sup>17</sup>, high implementation cost<sup>18</sup>, difficulties accommodating variations in an intervention (treatment heterogeneity), potential limitations on identifying causal mechanisms or reasons for observed impacts (unless coupled with other approaches) and ethical concerns. However, some of these concerns can be mitigated through good study design, or the addition of best-practice mixed-methods data collection. Moreover, most of these concerns are also present for quasi-experimental IEs that are much more commonly used to evaluate USAID programming, while in many cases it may be harder for quasi-experimental IEs to credibly overcome these challenges.

So, why are RCTs still uncommon in the development space? Some point to misunderstandings by IPs and programming decision-makers about how RCTs work (White 2013). In terms of potential ethical issues, concerns have also focused on a perceived unfairness about who receives an intervention or how randomization is determined, issues of informed consent for study participants, and power imbalances between evaluation researchers and the studied populations (Donovan 2018; Rodgers et. al. 2020; Hoffman 2020).

Skeptics have also raised concerns that the move towards RCTs has encouraged a disproportionate focus on smaller-scale and shorter-term interventions, because they tend to be more amenable to RCT approaches, at the expense of tackling evidence-based learning on bigger (and perhaps more important) development questions (Ravallion 2018; Donovan 2018; Bedecarrats, Guerin and Roubaud 2019; Rodgers, Bebbington and Boone 2020). Despite these concerns, RCTs are recognized to have played a seminal role in pushing forward development learning across a range of important development issues.<sup>19</sup>

Indeed, there is a growing middle-ground consensus that sees RCTs as an instrumental and powerful tool to learn about a range of important development questions and to inform policy-making on how to achieve poverty alleviation goals (Banerjee, Duflo and Kremer 2016; Rodgers, Bebbington and Boone 2020). In this view, RCTs are not the only useful learning tool available, nor are they always the best choice to learn about a given program or intervention, depending on a range of factors related to

---

<sup>17</sup> External validity refers to the extent to which one can reliably extrapolate results from one RCT context or intervention setting to another context or setting.

<sup>18</sup> But note that RCTs can often be less costly than using a quasi-experimental design to evaluate the same program, because quasi-experiments often require much larger sample sizes to obtain the same level of study power.

<sup>19</sup> This been discussed at length elsewhere. For a recent example, see Banerjee, Duflo and Kremer 2016.

context, intervention details, and so on.<sup>20</sup> RCTs are not always the only viable or strongest approach for all settings or problems at hand, and it is important to find the right fit between available evaluation and learning approaches and a particular intervention context (Ravallion 2020).<sup>21</sup> But, there are certainly many cases where an RCT approach can provide a more rigorous evidence base and added learning value compared to other available approaches, provided concerns about feasibility to implement can be overcome.

As RCTs have risen to prominence in the push for stronger evidence-based development learning, the World Bank, USAID and other donors have increasingly invested in using this approach to evaluate programming impacts (Bedecarrats, Guerin and Roubaud 2019).<sup>22</sup> The growing use of RCTs in the LRG sector over the past decade and the nature of ongoing debates around their use for development learning presents an opportune time to take stock of current LRG RCTs, and to particularly examine ethical concerns around their use in the sector with an eye towards summarizing lessons learned and best practices to mitigate such concerns.

## RCT CONTRIBUTIONS TO LEARNING ON LAND AND RESOURCE GOVERNANCE

### SUMMARY OF LAND SECTOR INTERVENTION RCTS

We identified 10 recently completed or ongoing RCTs of LRG programming through our literature search, and an additional seven LRG sector RCTs that are registered as a trial in development or ongoing, but for which we found no published or grey literature reporting outcomes to date.<sup>23</sup> We list these two groups separately in Tables 2 and 2a below, and focus most of our reporting on information from the 10 RCTs with some form of published results already available. This is because they contain additional learning on what the researchers experienced with respect to RCT implementation and potential challenges. This goes beyond the information available on the AEA registry, which is generally confined to details on the experimental design and basic study parameters.

A review of Table 2 provides some interesting insights regarding the current RCT evidence base for LRG interventions, with respect to the types of interventions covered, the geographic distribution, and basic elements of the RCT design. We see that most of the current RCTs focus on land formalization programs in the form of titling or provisioning of certificates that codify and recognize customary use rights to land. The distribution across major multi-lateral development organizations or foundations with a dedicated land portfolio or focus appears to be fairly even, generally at one to two studies each. Geographically, the RCT portfolio is heavily concentrated in sub-Saharan Africa. Across the 10 LRG interventions with a recently completed or ongoing RCT, eight are of interventions conducted in African

<sup>20</sup> This view spurred in part by the 2019 award of the Nobel Prize in Economic Sciences to development economists Abhijit Banerjee, Esther Duflo, and Michael Kremer for their application and popularization of experimental approaches, and particularly RCTs, to learning on poverty alleviation. For example, see multiple commentaries in the inaugural Symposium on Development and Poverty Alleviation launched in a 2020 World Development special issue (127: 104789).

<sup>21</sup> Indeed, this may also explain to some extent why RCTs to date have been more commonly applied in particular development sectors (such as health and education) and for certain types of interventions (such as cash transfers) (Sabet and Brown 2018).

<sup>22</sup> For example, USAID's Evaluation Policy states: "For impact evaluations, experimental methods generate the strongest evidence. Alternative methods should be utilized only when random assignment strategies are infeasible." (USAID, 2016).

<sup>23</sup> Of these, we further exclude one study from additional discussion in this report due to insufficient information in the AEA registry to enable complete characterization of the study, such as randomization approach, primary outcomes of interest and country of implementation. The remaining appear to be RCTs in progress for an additional seven LRG interventions in Bangladesh, Burkina Faso, Democratic Republic of Congo, Ghana, Philippines, Rwanda, and Uganda.

countries (Benin, Liberia, Mozambique, Tanzania (3 separate studies), Uganda, and Zambia), one is in Latin America (Bolivia) and one is in Asia (Mongolia). This geographic focus is even more skewed than the geographic distribution of development sector impact evaluations more broadly, where a third of development IEs are conducted in Sub-Saharan Africa<sup>24</sup> (Sabet and Brown 2018). In terms of basic RCT design, all of the 10 current RCTs employed a cluster-randomized approach and randomize treatment at the level of blocks, villages, or herder groups, while most of them measure household-level effects. Eight used stratified (block) randomization and two used a phased or pipeline design. Two of the RCTs employed a more complex multi-arm treatment design to examine individual and joint effects of complementary interventions.

Based on details available in the AEA registry, as seen in Table 2, we observe that the more recent generation of LRG RCTs currently under development are also heavily focused on sub-Saharan Africa (5 of 7). Four of the seven use a cluster-randomized design, and two of them employ a multi-arm treatment design. We also see some shifts in design approach, with perhaps more of a focus on smaller-scale studies to examine the impacts of more targeted aspects of interventions, rather than a general look at the average effects of land formalization interventions on a wide range of outcomes across many different outcome categories. Instead, these newer studies appear to be more narrowly focused on obtaining rigorous evidence about specific learning questions or causal mechanisms of interest, and they rely on innovative designs and randomization approaches to do so. If we contrast these with the RCTs in Table 2, we see some marked shifts in the “next generation” LRG RCTs. For example, more attention is paid to less studied types of interventions, and there seems to be a greater focus on how land programming can be coupled with other interventions to achieve intended gender equity and poverty reduction goals.

---

<sup>24</sup> Sabet and Brown 2018 report this as 34.4 percent.

**Table 2. Summary of LRG RCTs (Intervention Type and Key Characteristics of RCT Design).**

No	Intervention	Description of Intervention	Objective	RCT Focus and Key Findings	Donor	Country	Design	Unit of Assignment
1	<b>Land formalization</b> in urban unplanned settlements in Dar es Salaam  (Ayalew Ali et al. 2014; Ayalew Ali et al. 2016; Collin 2017)	Households were randomly assigned two vouchers (1. general and 2. conditional on including a woman as owner on title application) that could be redeemed for a discount when purchasing a formal land title.	Make titles more affordable to poor households in general, and improve the inclusion of women on titles.	<i>Focus:</i> Link between price incentives and gender inclusion; Do price incentives address obstacles to women's access to formal land ownership?; Peer effects on household propensity to adopt land certificate.  <i>Findings:</i> Households which were allocated a conditional voucher were much more likely to include a woman on their title application. However, it remains to be seen whether or not these strictly legal improvements in women's land ownership will result in actual de facto improvements in the lives of urban landowners.	World Bank	Tanzania	Conditional *	Blocks of land parcels
2	<b>Customary land formalization</b> in rural and secured the boundaries of all areas (establishment of parcels). Then, customary land ownership was formally and legally documented in the form of land use certificates.  (WB Gender Innovation Lab, 2019; Goldstein et al. 2015; Goldstein et al. 2018; Wren-Lewis et al. 2020.)	Communities identified, demarcated, and secured the boundaries of all areas (establishment of parcels). Then, customary land ownership was formally and legally documented in the form of land use certificates.	Improve tenure security and stimulate agricultural investment in rural areas.	<i>Focus:</i> Links between land demarcation and investment; Effects of land formalization on tenure security, investment, land transfers, agricultural production, women's empowerment; gender-differentiated effects.  <i>Findings:</i> Households in treatment villages were significantly more likely to report having parcels with clear borders. Improved land security following land demarcation led to increased long-term agricultural investment. Despite observed increases in investment, no average effects on agricultural output or farm yields. Also, no significant gender difference in impact on cultivation use, farm labor input intensity, etc. Later work (Wren-Lewis et al. 2020) found a positive impact on tree cover loss.	MCC	Benin	Conditional	Villages
3	<b>Community-based legal aid on land rights</b>  (Mueller et al. 2015)	The intervention allocated trained paralegals to villages. Any community members living in a treatment village could access to free legal advising and participation in legal education opportunities through the paralegal.	Identify the gendered impacts of access to legal aid on a range of land-related knowledge, attitudes, and practice outcomes.	<i>Focus:</i> Impacts of access to legal aid on a range of land-related knowledge, attitude, and practice outcomes.  <i>Findings:</i> Changes in legal knowledge, attitudes, and practices related to land are limited for	Unclear	Tanzania	Conditional *	Villages

		and, also, the cost-effectiveness.	women exposed to CBLA programs in the short term.				
4 Community land documentation in Uganda <i>(Knight et al. 2013a)</i>	Intervention consisted of four different legal services treatments: 1) full legal and technical support, 2) paralegal support and monthly legal education, 3) monthly legal education only, 4) control/minimal information dissemination.	Facilitate and support the protection of customarily-held lands by seeking formal documentation of community land claims.	<i>Focus:</i> Effects of land documentation on conflict resolution and prevention, governance, sustainable NRM, protection of land rights for vulnerable groups.  <i>Findings:</i> Community land protection efforts should combine the technical task of mapping and documenting community lands, the peace-building work of land conflict resolution, and the governance work of strengthening local land and natural resource management.	Gates Foundation	Uganda	Conditional	Communities
5 Community land documentation in Liberia <i>(Knight et al. 2013b)</i>	Intervention consisted of four different legal services treatments: 1) full legal and technical support, 2) paralegal support and monthly legal education, 3) monthly legal education only, 4) control/minimal information dissemination.	Facilitate and support the protection of customarily-held lands by seeking formal documentation of community land claims.	<i>Focus:</i> As above.  <i>Findings:</i> Community land protection efforts are not merely documentation exercises. New land laws and policies designed to protect communities' land and natural resources claims are urgently necessary.	Gates Foundation	Liberia	Conditional	Communities
6 Community land documentation in Mozambique <i>(Knight et al. 2014)</i>	Intervention consisted of four different legal services treatments: 1) full legal and technical support, 2) paralegal support and monthly legal education, 3) monthly legal education only, 4) control/minimal information dissemination.	Facilitate and support the protection of customarily-held lands by seeking formal documentation of community land claims.	<i>Focus:</i> As above.  <i>Findings:</i> Community land delimitation activities should combine the technical task of mapping and titling community lands with the peace-building work of land conflict resolution and the governance work of supporting communities to strengthen land and NRM and promote intra-community equity.	Gates Foundation	Mozambique	Conditional	Communities
7 Participatory mapping of village resources and land use on indigenous lands <i>(Victoria Reyes-Garcia, et al. 2012)</i>	Participatory mapping of villages within and around the Tsimane' Original Commutarian Lands. Additionally, the researchers produced a map representing villages' land and resource use, and conducted a communal workshop where they gave a copy of the map and explained its usefulness.	To examine the relation between participatory mapping and internal and external conflicts in rural communities, specifically the Tsimane'.	<i>Focus:</i> Effects of participatory mapping on conflicts.  <i>Findings:</i> Conducting participatory mapping did not produce any effect of real or statistical significance on either 1) the number of conflicts with outsiders entering Tsimane' villages, 2) the number of conflicts with Tsimane' from other villages, 3) negative attitudes or opinions of outsiders, or 4) negative attitudes or opinion of Tsimane' from other villages.	Unclear	Bolivia	Pipeline *	Villages

<b>8 Provisioning of pastureland leases to herder groups (IPA 2016)</b>	Drafting new legislation regarding rangeland and pasture use; mapping the rangeland; providing 15-year exclusive-use pastureland leases to groups of herder households; providing herder groups with infrastructure (e.g. materials for fences); and providing trainings in herd and pastureland management.	Improve the livelihoods of semi-nomadic herding households living in the areas surrounding Mongolia's larger cities.	<b>Focus:</b> Effects of  <b>Short-term (at midline):</b> Some evidence of reduced pasture load per hectare and control of herd size. No evidence of impacts on seasonal migration or livestock relocation patterns. No evidence of increased use of hay/fodder, or improved tenure security. <b>Long-term (at midline):</b> No evidence of improved land quality. Evidence of reduced animal mortality, but also decreased milk yield per milking cow.	MCC/ MCA	Mongolia	Conditional	Herder groups
<b>9 Certification of customary land rights (plus climate-smart agricultural support) (Huntington et al. 2018; Huntington and Shenoy 2021)</b>	Village-level participatory mapping, land administration support and mapping and provisioning of informal customary land use certificates for individual parcels; also facilitate tree planting adoption and survivorship on smallholder farms.	Strengthen customary tenure security, while also supporting agroforestry extension services.	<b>Focus:</b> Individual and joint effects of mapping, certification and agro-forestry support on tenure security, land governance, agroforestry uptake, agricultural investment and productivity, and livelihoods; gender-differentiated effects.	USAID (TGCC)	Zambia	Conditional	Villages
<b>10 Formalization of customary land rights (Persha and Patterson-Stein 2018; 2021)</b>	Assisted villages and local district land offices in completing village land use planning, mapping and registering customary use rights to individual parcels in 30 villages, and delivering Certificates of Customary Right of Occupancy (CCROs) to households. It also included education on land laws, CCROs, and land management.	Aim to increase land tenure security and lay the groundwork for sustainable agricultural investment.	<b>Focus:</b> Effects of land use planning, mapping and certification on tenure security, land management, disputes, land use and investment, women's empowerment, economic and environmental outcomes. Gender-differentiated effects.  <b>Findings:</b> The program had significant positive impacts on household tenure security and documentation of land rights, reduced the likelihood of current and future land disputes, and a smaller positive impact on use of communal land. No evidence for impacts on the likelihood of fallowing, crop diversification, household land investments, access to credit, or other indicators of household economic wellbeing during the evaluation timeframe.	USAID (LTA)	Tanzania	Pipeline	Villages

**Table 2a. Summary of “Next Generation” LRG RCTs (Intervention Type and Key Characteristics of RCT Design).**

No	Intervention	Description of Intervention	Objective	RCT Focus and Key Findings	Donor	Country	Design	Unit of Assignment
1	<b>Land titling coupled with a new credit line product</b> <i>(Ghatak, Maitreesh et al. 2019)</i>	First stage: villages randomly assigned to receive an offer for a fully subsidized land title. Second-stage: households within villages were randomly assigned to receive an offer for a credit line. (Intervention also includes gender conditionality and gender information treatments.)	To examine the standalone and complementary impacts of land titling and improved access to credit.	Ongoing	World Bank	Uganda	Cluster randomized; multi-treatment household (second stage)	Village (first stage); household (second stage)
2	<b>Effects of distribution of irrigated land, land preparation materials and inputs and land tenure documents via public lottery</b> <i>(Bambio, Yiriyibin and Christopher Ksoll 2018)</i>	Randomly distributed leases for 710 hectares of irrigated land in the Di perimeter, training in agricultural technologies for irrigated lands, starter kits (land preparation, materials and inputs), and land tenure documents.	Analyze the effect of winning the Di Lottery on agricultural practices, production, total agricultural income, overall household income and land tenure security.	Ongoing	MCC	Burkina Faso	Unclear	Unclear
3	<b>Effects of subdivision of collective land titles to formalize individual property rights</b> <i>(Gunnsteinsson, Snaebjorn 2020)</i>	Parcelization: subdividing land that is collectively titled and randomly issuing individual land titles to farmers previously listed under a collective title.	Understanding the impacts of rights on investment and the degree to which the lack of formal rights limits financial market development.	Ongoing	Unclear	Philippines	Cluster randomized	Groups of farmers
4	<b>Effects of provisioning of written contracts that specify land use rights and property and rental information</b> <i>(Karpe, Saahil et al. 2019)</i>	Two interventions: 1) introduce formal written contact, co-signed by village leader, that specify the property, dates of land use rights, etc.; and 2) asking lead farmer to serve as a “farmer broker”, and obtain information from other farmers to identify which farmers are interested in renting in or out land.	Test two interventions meant to resolve land market frictions in the context of potentially transformative technological change through irrigation in Rwanda.	Ongoing	Word Bank	Rwanda	Cluster randomized	Unclear

<b>5 Effects of access to irrigated land plots and land rental contracts</b> <i>(Goldstein et al. 2016)</i>	First stage: randomly allocates access to irrigated plots of land via rental agreement with Arikú Farms among interested households. Second stage: within each household, randomly allocate the contract to either the husband or the wife.	To provide evidence on whether providing direct access for women to productive inputs alters production or consumption patterns in the household, including whether or not aggregate output, productivity, and intra-household efficiency increases.	Ongoing	World Bank	Ghana	Stratified (block) randomization	Households (first stage); household member (second stage)
<b>6 Effects of performance scorecards on performance of bureaucrats</b> <i>(Mattsson, Martin 2018)</i>	Generate a monthly performance scorecard for bureaucrats in charge of land records.	Explore effect of performance scorecards on bureaucrat performance in charge of land records.	Ongoing	Unclear	Bangladesh	Stratified (block) randomization	Unclear
<b>7 Effects of providing subsidized access to formal property titles</b> <i>(Balan, Pablo et al. 2018)</i>	Three interventions: 1) providing citizens subsidized access to formal property titles; 2) collecting property taxes through local bureaucrats compared to centralized collectors deployed by the provincial ministry; and 3) offering property tax discounts of varying levels to households.	Explores three understudied determinants of low tax compliance: weak property rights, inefficient methods of tax collection, and liquidity constraints.	Ongoing	Unclear	DRC	Cluster-randomized	Polygon-level within the city of Kananga (taxation interventions); individual-level (property titling intervention)

## SUMMARY OF RCTS BY BROAD INTERVENTION TYPE AND OUTCOMES MEASURED

Table 3 below tallies the RCTs by broad intervention category, and distinguishes those categories that are the primary focus of the RCT (red dots) from other components of the intervention package. We categorize according to USAID's land sector and resource governance "roadmap" (Stevens et al., 2020), which classifies LRG interventions into five broad categories: (1) land use planning and natural resource management (NRM); (2) property rights and boundaries clarification, and official rights recognition; (3) capacity building of offices in land administration; (4) awareness raising of land rights and regulations; and (5) legal, regulatory, and policy dialogue, and advocacy and reform.<sup>25</sup> We note that more than half of the LRG RCTs we reviewed measure the cumulative impacts of several different interventions that were rolled out as part of the same program (for example, sensitization on land rights, combined with mapping and rights recognition, and support to a local government land administration system).

The focus of the current LRG RCTs (in terms of intervention type and impacts that the RCTs aim to measure), however, is split between an emphasis on effects of land rights recognition and effects of awareness-raising about land rights. Of the 10 interventions that comprise our core LRG RCT review, more than half investigated the impact of awareness-raising about land rights and regulations as either the main or a secondary focus of a property rights intervention. By far the most common awareness-raising interventions provided legal services to members living in the treated community. These legal services generally consisted of free legal advising on land rights via a trained paralegal, or providing regular education and trainings within the intervention timeframe. Property rights and boundary clarification interventions generally include programs that conduct land mapping and facilitate land rights recognition through some form of documentation to the landholder. Some of the studies also have some emphasis on land use and natural resource management. Few LRG RCTs to date have focused explicitly on impacts of land administration support (with the exception of the DRC RCT in Table 2a) or the effects of legal or regulatory dialogue and advocacy.<sup>26</sup> This makes sense intuitively, as these types of broader interventions, which are often implemented at national or a small number of high-level administrative units, may be less amenable to an RCT design. However, these gaps may also provide a window of opportunity for future innovative work.<sup>27</sup>

---

<sup>25</sup> Also see Lisher 2019 and Tseng et al. 2020 for additional categorization of LRG interventions.

<sup>26</sup> But see Sandefur and Siddiqi 2015 for an RCT of legal aid services in Liberia, including those focused on resolving land disputes, provided via community paralegals trained in mediation and legal advocacy.

<sup>27</sup> For example, RCT designs may be possible for administrative support or advocacy programs that are designed to be rolled out across several lower-level administrative units.

**Table 3. LRG RCTs by Broad Intervention Type.**

No.	Intervention under RCT study	Country	1. Land Use Planning & NRM	2. Property Rights & Boundaries Clarification, Official Rights Recognition	3. Land Administration Capacity Building	4. Awareness Raising of Land Rights & Regulations	5. Legal, Regulatory & Policy Dialogue, Advocacy, & Reform
1	<b>Land formalization</b> in urban unplanned settlements in Dar es Salaam (Ayalew Ali et al. 2014; Ayalew Ali et al. 2016, Collin 2017)	Tanzania		●	●	●	
2	<b>Customary land formalization</b> in rural areas (establishment of Rural Landholding Plans (PFR)) (WB Gender Innovation Lab, 2019; Goldstein et al. 2015; Goldstein et al. 2018.)	Benin		●	●	●	
3	<b>Community-based legal aid on land rights</b> (Mueller et al. 2015)	Tanzania				●	
4	<b>Community land documentation</b> in Uganda (Knight et al. 2013a)	Uganda				●	
5	<b>Community land documentation</b> in Liberia (Knight et al. 2013b)	Liberia				●	
6	<b>Community land documentation</b> in Mozambique (Knight et al. 2014)	Mozambique				●	
7	<b>Participatory mapping</b> of village resources and land use on indigenous lands (Reyes-Garcia et al. 2012)	Bolivia	●			●	
8	<b>Provisioning of pastureland leases</b> to herder groups (IPA 2016)	Mongolia	●	●		●	●
9	<b>Certification of customary land rights</b> (plus climate-smart agricultural support) (Huntington et al. 2018)	Zambia	●	●	●	●	
10	<b>Formalization of customary land rights</b> (Persha and Patterson-Stein 2021)	Tanzania	●	●	●	●	

● Denotes intervention categories that comprised the main focus of the RCT.

● Denotes interventions that were part of the overall intervention package, but the RCT does not appear to be designed to isolate impacts of this activity component(s) separately from other components of the intervention.

Next, we review the types of outcomes that LRG RCTs have focused on measuring. Table 4 below summarizes the outcomes examined by each RCT across broad outcome categories. Overall, shorter-term outcomes are more strongly represented across the LRG RCTs conducted to date. By far the most commonly measured outcomes in existing LRG RCTs are perceptions of tenure security (9 out of 10) and land disputes/conflict resolution (7 out of 10). Other common outcome measures include land documentation uptake (4 out of 10), agricultural or livestock investment and productivity (4 or 3 out of 10, respectively). Some studies have paid attention to gender differentiated impacts secondarily (4 out of 10), but most of these focused only on differences for female-headed households relative to male-headed households, and some did not appear to conduct formal hypothesis tests due to power limitations or other reasons. There is a notable lack of coverage on co-titling or specific measures of women's empowerment (such as women's bargaining power or intra-household decision-making). Other notable gaps include attention to the effects of LRG interventions on land rights knowledge, food security, land and labor markets and credit access.

**Table 4. LRG RCTs by Outcome Categories Measured.**

No.	Intervention	Outcome Category															
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	Land formalization in urban unplanned settlement in Dar es Salaam	●		●	●					●						●	
2	Customary land formalization in rural areas (establishment of (PFR))		●			●	●	●	●	●	●		●		●	●	
3	Community-based legal aid on land rights	●	●														
4	Community land documentation in Uganda	●	●	●													
5	Community land documentation in Liberia	●	●														
6	Community land documentation in Mozambique		●	●													
7	Participatory mapping of village resources and land use on indigenous lands			●													
8	Provisioning of pastureland leases to herder groups	●	●			●	●					●		●	●	●	
9	Certification of customary land rights (coupled with climate-smart agricultural support in some areas)		●	●	●		●	●	●	●			●	●	●	●	●
10	Formalization of customary land rights	●	●	●	●		●		●	●	●		●	●	●	●	●

No.	Outcome Category
1	Land rights knowledge and attitudes
2	Tenure security
3	Land disputes / conflict resolution
4	Land documentation uptake
5	Co-titling
6	Agricultural / livestock investment
7	Agricultural / livestock productivity
8	Land markets / transferability
9	Credit access
10	Women's bargaining power / intra-household decision-making / empowerment

No.	Outcome Category
11	Land allocation and administration
12	Food security
13	Social capital / collective action
14	Incomes / economic wellbeing
15	Labor market participation
16	Natural resource management / land use
17	Gender-differentiated effects

## **IV. PRACTICAL CHALLENGES, ETHICAL CONSIDERATIONS AND MITIGATING OPTIONS FOR IMPLEMENTING RCTs OF LRG INTERVENTIONS**

### **COMMON IMPLEMENTATION CHALLENGES FOR LRG RCTs AND MITIGATION OPTIONS**

Impact evaluations identify program impact by comparing outcomes between program beneficiaries to those of a control or comparison group of non-beneficiaries. The control or comparison group represents the counterfactual, or what would have happened in the absence of the program intervention. The key challenge for rigorous impact evaluations is producing valid estimates of the counterfactual. Other common design or implementation challenges for RCTs and other IE approaches include: dealing with selection bias<sup>28</sup>, threats to internal and external validity, imperfect compliance and low take-up rates for the intervention, treatment heterogeneity, contamination or spillovers,<sup>29</sup> sample attrition, attrition bias and other data collection challenges, and unintended behavioral effects from study participants (Winters et. al. 2010; Gertler et. al. 2016). Concerns about the randomization approach and randomized assignment of treatment, external validity, and unintended behavioral effects among the studied population may be more commonly experienced by RCTs relative to other IE approaches.

These general challenges are well-covered in impact evaluation toolkits.<sup>30</sup> Here, we draw on the published literature from land sector RCTs to highlight challenges that were experienced in practice and noted by the authors over the course of implementing land sector RCTs. Across the 14 RCT studies of land sector interventions that we reviewed, the four implementation challenges that researchers most commonly mentioned were:

- Concerns over delayed program implementation and the effects on evaluation integrity;
- Timing of endline data collection relative to when outcomes are expected;
- Spillover effects or contamination, whereby some members of the comparison group may have received at least some aspects of the intervention; and
- Imperfect compliance, in which the intervention that was received by some members of the treatment group diverged substantially from what was planned.

All of these situations are common potential challenges for impact evaluations that can affect confidence in and the validity of the impact results, although none of them are specific only to RCT designs. We highlight examples from the land sector and how they were dealt with below.

**Delayed or incomplete implementation** of the planned intervention was a common practical challenge experienced by the land sector RCTs we reviewed. Among the 10 interventions covered in our review, this occurred as a result of (1) slower than anticipated pace of government delivery of land documents or incomplete compliance with the original implementation plan by government or traditional authority counterparts (Huntington et. al. 2018; GIL 2019); (2) inability to proceed with the

---

<sup>28</sup> Selection bias refers to systematic differences between members of the treatment and comparison groups that are related to their outcomes under the intervention and, if not able to be addressed via the impact analysis, can lead to biased estimates of impacts. In theory, random assignment of treatment overcomes concerns on selection bias, but in practice it is possible for some randomization processes to generate unbalanced treatment and comparison groups either on observable or unobservable characteristics that also shape the outcomes under the intervention.

<sup>29</sup> Spillover effects are impacts of a given intervention, whether negative or positive, that occur to individuals or households who were not directly targeted by the program or included in the treatment group.

<sup>30</sup> For example, see Duflo et. al. 2007; Winters et. al. 2010; Glennester et. al. 2013; Gertler et. al. 2016.

intervention in some intended treatment areas due to government administrative decisions (for example see Ali et. al. 2014, where government re-zoning of land after the intervention started rendered some intended beneficiaries ineligible for treatment); (3) high transaction costs, which in one study constrained the ability of paralegals to advise and educate as planned in treated communities (Mueller et. al. 2015); and (4) context-specific, intra-community obstacles such as elites interfering with the land documentation process in a community (Knight et. al. 2013b; 2014).

In these cases, it is important for evaluation teams to be aware of such delays or changes to the planned intervention as they are happening, or at least well before the endline data collection. This implies close coordination between the evaluation team and program implementers, and frequent monitoring of treatment progress by the evaluation team throughout the intervention lifetime. In some cases, the evaluation and learning potential can be strengthened by adding questions to the endline data collection to better identify which study participants were affected by the delays, and how changes to implementation may have affected intended beneficiaries. Evaluation teams may also be able to partially mitigate the effects of such changes on evaluation integrity by oversampling at baseline so that the study remains well-powered even if some treatment units do not receive the intended intervention. Evaluation teams may also need to be flexible in when endline data collection will be conducted, to accommodate implementation delays if they occur.

**Timing of endline data collection** relative to when impacts are expected was one of the most commonly referenced challenges by RCT and other impact evaluations of land sector interventions. For six of the 10 interventions covered by the RCT studies we looked at, evaluation teams cited insufficient time between intervention activities and endline data collection as a potential explanation for null treatment effects. While donors are often interested to receive evaluation results as soon as possible, the timing of endline data collection must also be balanced against when short-term outcomes may be expected to accrue for beneficiaries after the intervention has been implemented. For LRG interventions, this timing must also often be somewhat flexible with respect to implementation progress, since LRG interventions can often face unexpected delays. This may be particularly the case for interventions that work with government partners to provide legal documentation, such as right of occupancy certificates or land titles, since the timing of document delivery also depends on government processes and timelines that are often outside the control of IPs. Three of the studies reviewed for this report noted delayed or incomplete provisioning of land certificates (Ayalew Ali et al., 2014) or land titles (WB Gender Innovation Lab 2019; Knight et al., 2013a) to intended beneficiaries. RCT evaluation teams for LRG programs must stay in close coordination with program implementers, and be somewhat flexible in when they might expect to conduct the follow-up round of data collection.

Evaluation teams implementing LRG-sector data collection face an additional challenge in timing endline data collection because several elements of LRG theories of change anticipate at least three to five year time frames for some key outcomes and medium-term impacts to be realized at scale. For example, results from Reyes-Garcia's (2012) study in the Bolivian Amazon indicated that the participatory mapping intervention under study did not produce any statistically significant effect on the number of conflicts, a key outcome of interest. The authors cited the short time between the participatory mapping and the post-intervention endline survey as a potential reason for their finding of null effects. In that case, the endline survey was conducted only one month after the completion of intervention activities. Huntington et al. (2018) also cited the short timeline between the close of the intervention and endline data collection (i.e., less than a year) as a potential reason for null effects of customary land

use certification on agroforestry investment and other indicators measuring land governance and agricultural productivity. They also noted that the absence of a treatment impact for indicators of longer-term impacts was expected, given the timing of the endline data collection. Since changes to shorter-term outcomes, such as perceived tenure security, may not accurately predict or correlate with longer-term results that practitioners hope to achieve (such as increased on-farm investments, or improved economic wellbeing), shorter-term outcomes also cannot on their own be used to infer longer-term positive impacts for a given LRG program.

Planning for multiple rounds of follow-up data collection after baseline is one way that donors and evaluation teams can mitigate this challenge, although this has implications for the overall cost of the evaluation. Also, the ability to track survey respondents after baseline may become more difficult with increasing time since baseline, resulting in higher sample attrition. Erring conservatively on estimates of sample attrition during the IE design stage, and oversampling at baseline, are two ways to help protect against this, although they also have cost implications for the evaluation.

Still, the number of land sector RCTs in our review that pointed to insufficient time between intervention end and endline data collection as a likely reason for null results on key outcomes suggest a need to carefully consider trade-offs between learning potential and a single round of endline data collection for evaluations of LRG programs. Conducting endline data collection shortly after an LRG intervention concludes, for example within a year or so post-intervention, is essential for capturing short-term effects and evidence-based learning within a reasonable timeframe for program decision-makers. But, where resources allow, planning from the start for an additional round of data collection an appropriate number of years later may be as essential for building a more comprehensive understanding of how LRG programs affect many core development objectives over the longer-term. The appropriate timing for such follow-on data collection will vary based on the intervention and longer-term impacts of interest, potentially ranging from five or more years post-intervention for many LRG interventions.

**Spillover effects (or contamination)** occur when an intervention affects non-participants. These effects can be positive or negative. If non-participants who experience a spillover are members of the control group, the control group no longer accurately represents what would have happened to the treatment group in the absence of treatment (i.e., the counterfactual). The impacts of development interventions can be substantially over- or under-estimated in the presence of negative or positive spillovers, if such indirect effects are not taken into account in the impact analyses. But, it is also possible to obtain valuable additional learning on program effects when spillovers are present. A key take-away for land sector RCTs is for evaluation teams to think through with IPs during IE design stage the potential channels for spillover effects from the intervention(s) to take place. Then, either plan for their possibility and appropriate measurement as part of the RCT design or, where possible, take additional steps to minimize the possibility for spillover to reach the control group. If logic suggests the potential for spillovers will still be high despite those efforts, aim for an evaluation design, sample size and sampling strategy that will enable testing for spillover effects where possible.

One such example from LRG RCTs was presented by Mueller et al. (2015), who examined the impacts of gender-differentiated access to legal aid on a range of land-related knowledge, attitude, and practice outcomes in northwestern Tanzania. Seventy villages were randomly assigned to receive trained paralegals, from whom any village member could access free legal advising and participate in legal education opportunities. An additional 69 villages were assigned to the control group. Mueller et al.

cited two sources of possible information spillover from the intervention to the control group in their study context: first, the close proximity between some treatment and control villages could make it easier for individuals from the treatment group to know or come in contact with comparison group individuals, and potentially share information. Second, in the region of study, religious places of worship could service multiple communities and serve as an inadvertent source of information sharing across members of the treatment and comparison groups. In this case, the evaluation team found evidence of paralegal exposure in both the treatment and control villages, with men in control group villages that were closer to the intervention villages being more aware of the paralegals than control men overall. In such situations, evaluation teams can often use the variation in spillovers across different control group units to obtain additional learning about the intervention effects, if the sample has been designed with this purpose in mind.

Similarly, in a 2013 evaluation of a community land titling initiative in Rivercess Country, Liberia, researchers found that close proximity of treatment and control group communities facilitated spillovers, and members of the control group communities frequently obtained support from the paralegals (Knight et al., 2013b). Although access to the paralegal support may have positively impacted these control communities, if the spillover cannot be addressed through the impact analysis (as was the case for this intervention), it can result in a distorted estimate of the effect of the intervention. On the other hand, if the IE design can plan for the possibility of spillovers from the outset, and construct an appropriate and sufficiently powered sample to accommodate its measurement, the evaluation can generate additional learning about unintended consequences of the intervention, whether positive or negative, for those who were not directly planned to benefit from the intervention. In practice, this will often mean increasing the sample size (and budget) for the evaluation in order to obtain sufficient power to accommodate the testing for spillover effects.

A proactive way to approach this in LRG RCTs is for evaluation teams and IPs to think through possibilities for unintended spillovers together during the IE design phase, and seek to design the IE such that it will be possible to measure intervention effects in the presence of spillovers if this seems likely. In practice, often this means at a minimum that the evaluation sample size must be large enough, the data collection tools are designed to collect information on potential treatment exposure in the control group and secondary information is also collected on geographic proximity of intervention and comparison units, social networks, and other potential channels by which information or other benefits from an intervention could reach non-participants and members of a comparison group.

In terms of learning value-add, while spillovers are not always desirable from an evaluation standpoint, they can in some situations be exploited for additional learning, such as to estimate the effects of different treatment intensities. For example, methods are available to measure not just the average treatment effects for those who received an intervention, but also the effects on an ineligible population in the presence of spillovers. Spillovers can also provide important learning on the mechanisms by which an intervention operates. Accounting for them in an IE design can thus help ensure that policy recommendations coming out of the evaluation are based on more accurate estimates of effects, and an understanding of who else might benefit or be negatively impacted by the intervention outside the intended treatment group (Angelucci and Di Maro 2015).

**Imperfect compliance** occurs when there is a difference between the assigned treatment status as planned by the intervention at program start, and the actual treatment status once implementation gets

underway. Imperfect compliance can manifest in a variety of ways, but often has substantial implications for the integrity of the planned IE. For example, some intended participants may be excluded from the program because of administrative or implementation errors, or changes beyond the control of the intervention. A recent example from the land sector comes from a planned RCT of a land titling intervention in Tanzania (Ali et al. 2014). In two communities in Dar es Salaam, households in treatment areas were assigned vouchers as an incentive to purchase a formal land title. However, excessive flooding prompted government authorities to re-zone part of the treatment area after the intervention began, such that those residents were no longer eligible for land titling. As a result, overall participation and take up of the land titling intervention was significantly decreased, with substantial implications for the power of the impact analysis to detect changes to outcomes at the levels that were anticipated at evaluation design, due to changes in sample size for the treatment and comparison groups. In such situations, the learning opportunities from the evaluation are often diminished.

Another example comes from the World Bank's Gender Innovation Lab's impact evaluation of land formalization interventions on the investment and welfare of households in Benin (GIL 2019). The program contained two interventions: land demarcation activities and the delivery of legally valid and transferable land use certificates to individual landholders. The issuance and delivery of these certificates was led by local governments and village land committees and was carried out at a much slower pace than the first intervention. As a result, only 19% of the demarcated panels had received a Rural Landholding Certificate by the time of the second survey wave, with implications for the viability of impact analysis at that stage. In a third example, from USAID's RCT of the Tenure and Global Climate Change program in Zambia, one of the local chiefs working with the program to provide customary land certificates chose to hold off on authorizing the distribution of the certificates to constituents in the treatment villages until communities in the control group had also been treated (a decision that apparently stemmed from his own ethical concerns about some members of the chiefdom not being able to receive the service). Although contrary to the initial implementation plan, it stemmed from a desire for all constituents in the chiefdom to benefit from the program. This reduced the fully treated sample for the evaluation at the time of endline data collection, and created a new category of partially treated beneficiaries, but the evaluation team was able to incorporate some ways to mitigate this through different analyses conducted at endline.

Another variation in imperfect compliance occurs when not all of the intended program participants ultimately decide to participate in the program. This can especially occur for land sector interventions where the intervention is available to all members of an eligible group, and individuals, households or communities within that group can decide whether or not they would like to participate (for example, a situation where land titles are universally available to landowners through the program, but individuals can choose whether to receive one). Such was the case for a community land documentation program in Uganda (Knight et al. 2013). Although communities had volunteered to participate, they began to withdraw soon after the study began. Per the study authors, the most common cause of this withdrawal in the study context was the influence of community elites or powerful individuals who – fearing the loss of lands they had appropriated in bad faith – took steps to raise opposition to the project. This threatened to compromise the validity of the RCT. In response, the implementer added 14 additional communities to the treatment group slated to receive the program intervention.<sup>31</sup> Lower than expected

---

<sup>31</sup> Of course, this type of development and the ability for community elites to influence the process in itself also provides important learning for LRG programming efforts to document community land (Knight et al., 2013). It also underscores the importance of gaining a nuanced understanding of the political economy of an intervention area and related context early in

take-up of LRG programming is something that can also be planned for at least to some extent during IE design phase, and can often be at least partially mitigated through oversampling at baseline and being conservative about take-up expectations in the power analysis at IE design. Similarly, if members of an eligible group can choose to self-select into participating in the program, then the evaluation team may also need to account for this source of selection bias in the IE design and analysis.

**TABLE 5. COMMON IMPLEMENTATION CHALLENGES FOR LAND AND RESOURCE GOVERNANCE RCTs**

KEY IMPLEMENTATION CHALLENGES	EXAMPLES FROM PUBLISHED STUDIES	NUMBER (N = 10)	POTENTIAL MITIGATION STRATEGIES
<b>Delayed or incomplete program implementation</b>	Policy decisions by government led to a delay in the production of titles for treatment communities. (Ayalew Ali et al. 2014; 2016)  High transaction costs constrained the mobility of paralegals in treated communities (Mueller et al. 2015)  Local elites influenced community interest to participate in land documentation process (Knight 2013b; Knight 2014)	6	Maintain close coordination with IPs throughout activity lifetime; Plan for flexibility on when the endline round of evaluation data collection takes place; Consider oversampling at baseline.
<b>Timing of evaluation endline data collection</b> relative to when short and longer-term outcomes and impacts are expected	Timing between the intervention and the endline survey round may have been too short to capture the effects of participatory mapping on internal and external conflicts.  EXAMPLE: Reyes-Garcia (2012)	6	Plan from design phase to conduct multiple follow-up rounds of data collection.
<b>Spillover effects (contamination)</b>  Intervention affects non-beneficiaries, including members of the control group	Individuals in control villages were exposed to paralegals, due the close proximity between treatment and control villages.  EXAMPLE: Mueller et al. (2015)	3	Conduct logic exercise on the potential for spillovers and possible channels at design phase; structure the evaluation sample and data collection instruments to enable measurement of spillovers.
<b>Treatment non-compliance</b>  Units assigned to the treatment group do not receive treatment	Communities that had volunteered to participate ended up withdrawing after powerful elites engendered opposition to the program within their communities.  EXAMPLE: Knight et al. (2013b)	3	Plan for additional units in the treatment group from design phase; Maintain close coordination with IPs throughout activity lifetime; if appropriate, support M&E data collection by IPs during activity lifetime to understand reasons for non-compliance; structure follow-up data collection to enable additional learning on unintended effects.

In addition to these four issues that were more commonly noted across the studies we reviewed, some other practical challenges were also encountered and noted in these published studies. These challenges are also instructive for land sector RCT design and future planning more generally. They include:

- Delays in receipt of government information needed to finalize the randomization process (Ali et. al. 2014 and 2016; Collin 2017);
- Lower than anticipated take-up rates for the intervention (Ali et. al. 2014; IPA 2016);
- Unanticipated costs to beneficiaries to receive land documents that affected intervention roll-out and delivery of the documents (GIL 2019);
- Challenges isolating the effects of particular intervention components of interest due to concurrent timing of different elements of an intervention package (Goldstein et al. 2015, 2018);

---

project planning stages (for example, during project start-up), by IPs and evaluation teams alike, to potentially help anticipate and plan strategically for such challenges.

- Randomized assignment of units into treatment and control groups did not achieve its intended aim, resulting in a lack of balance across key characteristics of the treatment and control groups (Goldstein et al. 2015, 2018)

In addition, many RCTs fail to consider potential spatial bias in the sample design, which can result in treatment and control group assignment that is not truly randomized or balanced on underlying characteristics that can affect outcomes. Cluster-randomized RCT designs can be particularly vulnerable to this, as they often draw on geographic areas (such as villages), as the unit of randomization. While delving into this issue in detail is beyond the scope of this report, we note that studies increasingly call attention to a need for RCTs (and IEs of all designs) to make better use of spatial analysis methods to ensure that findings are not vulnerable to hidden biases due to spatial effects and/or the spatial location and distribution of units for the study.<sup>32</sup>

### **ETHICS IN LAND AND RESOURCE GOVERNANCE RCTs**

Ethical concerns around the use of RCTs in the development space have been discussed at length, but here we take a closer look at these issues in the context of RCTs to evaluate LRG interventions. USAID's own guidance on impact evaluation brings up two key ethical critiques of RCTs: (1) RCTs may deny services to potential beneficiaries; and (2) For interventions that aim to target participants who are most in need, targeting requirements may need to be relaxed in order to obtain a large enough sample for the evaluation, such that some participants may not actually be among the “neediest” category. For the first issue, this could mean that those who do not receive the LRG intervention continue to experience harms due to unresolved land disputes, a greater sense of tenure insecurity and so on, while those who benefit from the intervention may see improvements on these issues.

However, mitigation strategies are often available to overcome both of these concerns. For the first critique above, a phased or pipeline RCT design, in which the intervention plans to provide the same service(s) or intervention to the control group after the evaluation is conducted, is one common mitigation strategy that can be planned for from early in the intervention or evaluation design. USAID also highlights that there is an ethical argument to be made about demonstrating an intervention's positive effects (or, at least not negative), particularly if the intervention is innovative or as yet untested. The benefit of doing so from a rigorous RCT may outweigh concerns about denying services to a comparison group during the lifetime of the evaluation. For the second critique, regression discontinuity designs may also provide a viable alternative in some situations, in which the treatment and comparison group for the evaluation are constructed around an eligibility cut-off to receive the program.<sup>33</sup>

Key ethical concerns for development RCTs often focus on issues of targeting and the vulnerability of potential beneficiaries. While concerns are often voiced over withholding potentially positive benefits from an eligible group of people who were arbitrarily selected as the control group, a phased RCT design is a viable way to overcome this concern. In the context of humanitarian interventions, it is also recommended that decisions on whether to have a pure control group should take into account how

---

<sup>32</sup> For additional reading on this issue, see for example: Bayliss, K. and A. Ham, 2015. How important is spatial correlation in randomized controlled trials? (No. 330-2016-13548). Available at: <https://ageconsearch.umn.edu/record/205586/>

<sup>33</sup> For additional discussion on regression discontinuity designs, see:

[https://www.usaid.gov/sites/default/files/documents/1870/IE\\_Technical\\_Note\\_2013\\_0903\\_Final.pdf](https://www.usaid.gov/sites/default/files/documents/1870/IE_Technical_Note_2013_0903_Final.pdf)

feasible it would be for the implementing organization to accommodate this in their intervention roll-out, together with a consideration of the level of vulnerability of the target population (Hoffman et al. 2020). There are also situations where the randomized selection of who should receive an intervention can in some situations actually be seen as the fairest approach (Dalziel 2017). This may particularly apply to situations where the program resources are limited and cannot be applied to all individuals, villages, or other beneficiary units that might like to receive the development intervention or service provided.

Ultimately, decisions on fairness or related potential ethical concerns for a given RCT will depend on the type of intervention under consideration, the options available for how the randomized selection may be conducted, and the beneficiary and implementation context. In some situations, a randomized approach may not be possible. In one LRG example, program implementers, USAID and the evaluation team were prepared to implement an RCT of a land titling intervention using a phased roll-out design, in which all of the eligible municipalities would eventually receive the intervention. However, an RCT approach was ultimately determined to be infeasible due to implementation concerns over meeting the timing of the phase-in schedule, government priorities on where the intervention should take place and concerns that it would not be possible to communicate about and conduct the randomized selection in a way that would be perceived as fair and transparent across a large number of eligible municipalities. As randomization at lower levels within selected municipalities was also determined to be infeasible, the evaluation team was able to pivot to an alternative IE design instead.

A key lesson for potential RCTs of LRG interventions is for USAID or the relevant donor, implementing partners, the evaluation team, and relevant government or other in-country counterparts to be well-informed on the potential options, trade-offs and benefits of randomized approaches, and have opportunities to discuss and work together on a feasible way forward from early in the intervention planning and IE design. This is often an iterative process that requires multiple discussions and information-sharing among the partners, so evaluation teams and USAID partners should also plan to build in sufficient time for that process to take place.

**Two USAID-supported RCTs of land sector programming, in Tanzania and Zambia, were able to accomplish randomized selection of villages to receive village-wide customary land certification programs by working closely with local government and customary authorities from early stages of planning.**

*In both cases, project implementers together with the evaluation team worked with local authorities to introduce the intervention objectives, benefits of an RCT evaluation to learn about impacts, and draw up a list of eligible villages. The teams then conducted the randomized selection of villages that would receive the program and villages that would serve as the control group. To do so required close coordination among the IPs, evaluation team and local authorities, and careful planning and sensitization with local authorities as to the benefits of this approach. For each program, all households within the selected treatment villages were eligible to receive the planned tenure strengthening services. Neither of the two programs were designed at the outset to provide the same services to the control group. However in both cases, as evidence on benefits of the interventions became available, USAID was able to extend the program activities to the control villages through follow-on programming.*

*USAID was also able to leverage the RCT evaluation design and follow-on programming to conduct additional evidence-based learning on longer-term impacts of the LRG interventions and targeted follow-up studies of issues that were highlighted through the earlier RCT results.*

*The approach did entail some trade-offs. From an implementation perspective, IPs noted that following the planned randomization approach and schedule imposed constraints on how they planned to carry out the interventions. This included introducing some inefficiencies to the timing and logistics of their work across different villages. In the Tanzanian case, the evaluation team worked with IPs to try to minimize those constraints to the extent possible, while frequent communication and coordination among IPs, the evaluation team and USAID helped to identify opportunities for problem-solving and alternative solutions. In both cases, USAID's facilitation of that communication and coordination was seen as instrumental for smoothing the process to the extent possible, and enabling innovative LRG programming and rigorous evidence-based learning from that programming to proceed simultaneously.*

## RCT DESIGN CONSIDERATIONS

One of the most basic and well-documented objections to RCTs in the development space is discomfort with the idea of experimenting on people (White, 2013; Singer et al., 2019): Is it ethical to allocate treatment to some individuals and withhold treatment from others? Best practice ways to address this concern include transparent communication with programs IPs and related stakeholders (local authorities, higher-level government administrators, USAID Mission and other OU staff, and potentially community leaders and members themselves depending on the level at which randomization is being conducted), about the reasons for randomization, the potential benefits, and the randomization options that might be available. It is also important for evaluation teams to provide sufficient information and seek buy-in about the potential for an RCT design from early in the evaluation design process. Then, there is a need to maintain close involvement of these stakeholders throughout the planning. USAID can

play an essential role in this by helping to facilitate this communication from early in the evaluation design phase, recognizing and allowing for the additional time and cost that go along with this sensitization and communication process, and by clarifying for all USAID's main learning priorities out of the evaluation.

The timing of RCT design phase activities is also very important. Evaluation teams must be able to hold these discussions with IPs while IPs are still in the process of designing their intervention and finalizing implementation details. This allows ample time for both teams to discuss if and how randomization might be feasible, the learning issues that IPs may be interested in from an independent evaluation, and how particular intervention details or aspects of implementation can be planned to meet those objectives.

If feasible, discussions may eventually also include potential beneficiaries of the intervention as well, for example once the RCT approach and randomization strategy has been determined. But while many of the LRG RCTs we reviewed appeared to have worked closely with implementers or governments to conduct the randomized allocation of beneficiary units into treatment and control groups, direct involvement of potential beneficiaries in the process was not common. This is likely due to the nature of the interventions and the unit at which randomization typically took place. Where we have seen involvement of potential beneficiaries in the randomization process, typically this has been in situations where a public lottery will be held in one or a small number of villages to select recipients for a very discrete intervention such as a farmer training, a voucher for a particular good or service, or a tree-planting contract.<sup>34</sup>

Once overall eligibility for the intervention has been established, LRG RCTs must also determine the options for how program benefits will be assigned to participants. This involves defining the **unit of assignment** for randomizing who will receive the intervention. LRG interventions can be applied at several different levels, depending on the nature of the intervention. For example, an LRG intervention can be provided to individuals or households, or it can be rolled out at a more aggregate levels, such as to neighborhood blocks, communities, municipalities, or larger administrative units. Due to a combination of the nature of the interventions and logistical and ethical viability, many RCTs in the land sector are designed as 'cluster-randomized' RCTs, in which randomization into treatment or control groups is done at the level of a cluster, such as a village or neighborhood. The treatment itself is then provided to all eligible sub-units within the cluster, such as households or individuals. All of the 14 LRG RCT studies we reviewed employed this type of cluster-randomized RCT design, with the unit of assignment either as blocks of land parcels, villages, communities, or herder groups.

It is also possible to randomize at the level of individual households for LRG interventions, but this is often seen as inivable due to concerns about spillover effects among neighbors, friends or individuals within the same network within a given village, overly high administrative burden on implementers to implement in this way, and ethical concerns related to fairness of some people in a village receiving land or resource governance services while others in the same village do not. However, we also note that some of the "next generation" RCTs in our review do appear to be implementing household-level or sub-village randomization of the intervention, in contrast to prior studies. These RCT studies also appear to be relatively smaller in scale and designed to provide evidence about impacts for a targeted subset of issues related to a broader intervention, such as how to increase co-titling rates, rather than

---

<sup>34</sup> An exception from the land sector is a public lottery that was held to select recipients of irrigated land plots among a group of eligible farmers, as part of an MCC-supported RCT evaluation of an agricultural development and land tenure program. (See: <https://assets.mcc.gov/content/uploads/evalbrief-20200022501-bfa-diverse-agriculture.pdf>)

all impacts from the intervention across a wide range of issues. In such cases, where the RCT is looking at differences in outcomes across intervention approach, randomization of who receives the particular permutation under study across households or individuals within communities is likely more feasible to achieve.

While cluster-randomized RCT designs have pros and cons, as with any IE approach, they are generally feasible to implement for many types of LRG interventions, provided there is sufficient buy-in, and may also entail lower data collection costs and implementer burden than other RCT approaches. A majority of the studies we reviewed evaluated community land titling and official rights documentation interventions for members of those communities. Here, the intervention is typically rolled at the community-level in any case and is available to all members within a community, hence a cluster-randomized design for the RCT may be the most logical choice in any case.

Still, being mindful of ethical critiques of RCTs around power asymmetries across researchers and implementers on one hand, and developing country participants on the other (Singer et. al. 2019), evaluation teams, donors, and implementers in the LRG sector may also need to take into account the process by which communities that are selected to receive LRG interventions have been defined, and who had agency to determine that process. Otherwise, as a recent LRG RCT on community land documentation experienced, such work could risk reinforcing long-standing inequities or historical norms that may themselves have either been externally imposed on intended beneficiaries, or end up reinforcing inequitable gender norms or other power imbalances within communities.

**Three RCTs of community land documentation programs encountered an important issue with respect to community-level implementation: who gets to define the socio-political and geo-spatial boundaries of the clustered unit, and how appropriate is the resulting definition?**

Knight et al. (2013a, 2013b, 2014) evaluated community land documentation programs in Uganda, Liberia, and Mozambique. In Northern Uganda, common grazing lands are central to village life. In response to growing population density, increasing land scarcity, and other pressures, the Land and Equity Movement in Uganda (LEMU) and the International Development Law Organization (IDLO) implemented an intervention to support communities to document and protect their customary land claims. In practice, however, the “community” unit of assignment – as defined by administrative state-drawn boundaries – did not neatly correspond to customary boundaries as understood by the multiple villages that shared communal grazing land.

This required the implementing team to conduct additional work to understand community boundaries on the ground and adjust its strategy to ensure all villages with ownership and use rights were included in the land documentation process. Additionally, there were cases in which the program was not implemented because one village within a government-delineated community rejected the intervention while remaining villages wanted to move forward with documenting their lands. Knight et al. (2013a) concluded that, rather than accepting externally imposed definitions of communities, LRG implementers should first work to understand how accepted these delineations are in practice, and, where needed, support communities to define their own social units and spatial delineations through extensive and highly participatory discussions before proceeding with community land documentation. While the feasibility of this may vary across different contexts in practice, evaluation teams should also be part of such discussions as changes to community boundaries or definitions will also have implications for evaluation design and implementation.

In addition to determining the rules for treatment eligibility and allocation, LRG RCTs must also determine the **method that will be used to randomize units into the intervention and control groups**. In a classic randomization design, all eligible units are randomly assigned to either treatment or control groups. This is often considered the most equitable method of randomization because all eligible units have an equal chance of receiving program benefits.<sup>35</sup> However, an argument can still be made that under this method the treatment is being withheld from some portion of equally eligible, equally deserving (or potentially more deserving) units.

If resources allow, an alternative approach is to use a **pipeline, randomized roll-out or phase-in randomization design**, in which all eligible units receive the program, but at different times. Instead of randomizing which units will receive treatment, the program instead randomizes the order in which

<sup>35</sup> A permutation of this approach in some development programming contexts is to conduct a public lottery, in which potential beneficiaries are also part of the selection process. The feasibility of this approach depends on the nature of the intervention and the randomization unit, and to date appears less viable for many village-level land sector interventions. In other sectors, public lotteries have been held within villages, for example to select individual farmers or households who will receive a given service, technology, voucher, or similar type of intervention.

each of the units received treatment. Units that are selected to serve as the control group will eventually receive the intervention as well, during a subsequent phase of the program. After baseline, the data collection for the RCT is then timed to correspond to the different phases of roll-out of the intervention across the one or more subsequent groups of control units. IPs must work with the evaluation team and other implementing stakeholders to decide if and how information will be shared with the control units on program roll-out. If members of the control group know that they will also eventually receive the intervention, they may change their behavior in the short-term in anticipation of the eventual program, and this could be in ways that affect the impact estimates.

It is also important for IPs and evaluation teams to work out the timing between phases and evaluation data collection collaboratively, and adhere to that schedule over the course of implementation so that the evaluation is able to measure both short-term and longer-term effects. If the time length between phases is too short, there is a risk the RCT may underestimate impacts or produce otherwise unreliable estimates of impact, because there was insufficient time for effects to accrue. The design must ensure sufficient time between the first and subsequent rounds of treatment to detect impacts from the intervention. Depending on how the roll-out is structured, the data collection for this type of RCT may require additional rounds beyond a simple baseline and endline.

Another option is to provide either some or all program benefits to control groups after the implementation period for the initial intervention has ended. For example, researchers in the Bolivian Amazon used the same treatment protocol to conduct participatory mapping with villages in the control group after the post-intervention survey had been carried out (Reyes-Garcia et al., 2012). Similarly, USAID extended its Land Tenure Assistance (LTA) activity in Tanzania to all eligible villages in the evaluation control group after the intervention had been completed in the treatment villages and the endline data collection for the RCT had been conducted. By this time, district and higher-level government authorities had seen positive effects of the intervention and were highly supportive of rolling it out in additional villages. Similarly, some of the control group villages that were adjacent to treatment group villages had become aware of the land formalization services provided to the treatment villages, and expressed interest to participate. A potential drawback for that example, however, stems from changes IPs made to the nature of implementation for this additional roll-out, whereby villagers were required to pay a small sum to receive land documentation that the initial treatment group received at no cost. The intent was to improve the sustainability of the service provisioning beyond donor support, hence expand the ability in future for more potential beneficiaries to receive the services. This type of modification may also provide an opportunity for valuable learning for LRG programming more generally, though there may also be a risk of perceived unfairness among the latter group of beneficiaries.<sup>36</sup>

### ADDRESSING POTENTIAL RISKS AND HARMS

The last overarching ethical consideration that warrants general discussion is the possibility that program beneficiaries and/or non-participants may be harmed by the RCT itself. In other words, even if evaluation teams are working on the assumption that the programs they are evaluating will be positive

---

<sup>36</sup> We note, however, that similar modifications and potential risks are incurred through many efforts to scale-up development programs, where the nature of implementation is necessarily modified to test options for improved sustainability of service provisioning across an expanded beneficiary base. The potential for the latter group of beneficiaries to learn about and feel unfairness over such differences is not unique to RCTs.

or, at the very least, neutral, they must also consider whether it is possible for unanticipated negative consequences to occur as a result of the randomized evaluation design. The likelihood of this in practice is expected to vary depending on the nature of the intervention and beneficiary context, and how randomization is conducted. It is therefore recommended that evaluation teams take this possibility into consideration from the earliest stages of IE design in collaboration with IPs, and work with IPs to monitor potential harms to beneficiaries or control group study participants throughout program implementation. None of the 14 LRG RCT studies we reviewed included discussion on potential risks to program participants and/or nonparticipants in their published papers as a result of the RCT, although it is likely many of them were required to address this while seeking institutional approvals to conduct research on human subjects.

Evaluation teams and donors must also consider the possibility that the control group could be made worse off or prevented from accessing otherwise available services, if not for the RCT. The answer to this will also depend on the nature of the intervention and the RCT design. However, we note that it is uncommon for RCTs to deny services to a control group outright. Just as members of the intervention are typically offered a given treatment and can then choose themselves whether to participate, members of the control are not offered the treatment but they also are not excluded from receiving it. An example from the land sector is USAID's RCT of customary land certification in Tanzania through the Land Tenure Assistance activity. All members of the 30 randomly selected intervention villages were offered customary land mapping and registration services through a systematic village-wide process, and households in those villages could then choose whether to participate or not. Members of the 30 randomly selected control group villages were not offered this service, but they could still request and pay for customary land mapping and registration through local authorities via the same government process that was already available to all households in the study area prior to the RCT.

With respect to potential harms or negative effects stemming from the LRG intervention itself (rather than from the RCT to measure impacts of the intervention), two of the LRG RCT studies we reviewed did explicitly discuss potential harms or unintended negative consequences to beneficiaries as a result of the program itself. In one case, as a result of the Land and Equity Movement in Uganda (LEMU) facilitating formal documentation of community land claims, Knight et al. (2013) found that in communities with a high-degree of internal dysfunction, the intervention's provisioning of external legal and technical support to community members had at times exacerbated intra-community conflicts. In other settings, there has been speculation that some LRG interventions, such as customary land formalization, could increase women's risk of gender-based violence – for example if strengthening women's land knowledge and rights might expose them to unanticipated repercussions among male family members.

LRG evaluations of USAID programs often seek to address unanticipated positive or negative effects of the intervention through a combination of quantitative survey data collection and qualitative data collection explicitly designed to explore these possibilities across different sub-groups of interest. While this issue is not unique to RCT designs, we include it here to call attention to a need for LRG RCTs to also maintain visibility on unexpected and possible negative consequences of the interventions under study, to the extent possible. Often this can be more comprehensively done through a combination of mixed-methods approaches that go beyond impact estimates derived from quantitative household surveys and statistical analysis alone. As evidence of programming effects from RCTs are often used to make decisions about program scale-up, replication, set aside or broader programmatic decision-making,

it is worth noting that evidence of unintended harms is just as important as evidence of positive impacts. Here, it is possible that an RCT may be better situated than other evaluation approaches to provide credible evidence on such unintended negative effects, given the higher level of rigor and confidence in results that RCTs often provide relative to quasi-experimental IEs or performance evaluation designs.

### COST-EFFECTIVENESS ANALYSIS AS AN ADD-ON CONSIDERATION FOR LRG RCTs

Despite the clear gains made to development learning around causal effects of LRG interventions through RCTs and rigorous quasi-experiments, to date there has also been a noticeable lack of attention to the monetary costs and the cost-effectiveness of such interventions in relation to the benefits and impacts achieved. This is not unique to LRG interventions within the development space. In their review of 2,000 agricultural impact evaluations, Mogues et al. (2019) found that fewer than 5% included any meaningful cost data. Of the 14 LRG studies we reviewed, only three explicitly addressed the need for cost-effectiveness information (Mueller et al., 2015; Knight et al., 2014; Mogues et al., 2019) and only one (Mogues et al., 2019) actually conducted a complementary cost-effectiveness analysis (CEA). This could be a missed opportunity for LRG RCTs, in terms of the overall contribution and value-add of such studies to the evidence base, as cost-effectiveness information is generally desirable to donors and could help increase the utilization and perceived value of RCT results for broader programming decisions.

Understanding the costs needed to achieve a given level of benefits is crucial for drawing policy and programmatic conclusions and for informing the allocation of scarce public resources, especially given the typically high price tag of land sector interventions rolled out at scale. Some have argued that any policy relevant impact evaluation should always go beyond measuring impacts to also unpack why a program does or does not work in a particular setting and whether it should be taken to scale or replicated elsewhere (White 2013). Many LRG IEs address such issues through mixed-methods evaluation designs that employ qualitative data collection to help understand reasons for impacts, and how and why different sub-groups of beneficiaries may have been affected differently by the intervention.

While attention to those issues is increasingly common in LRG RCTs and other IEs, attention to cost-effectiveness has remained a gap in the growing knowledge base. This is also worth considering as part of a broader discussion on RCT ethics, as one could also go so far as to view a given intervention as unethical even if it does work, if it is much less cost-effective than a viable alternative, since it ultimately will benefit fewer people or result in smaller impacts for the same amount of resources (Singer 2019).

Mogues et. al. (2019) summarized the overall benefits to performing a cost-effectiveness analysis of a 2010-2013 agricultural production advisory program implemented in three provinces of Mozambique. With the aim of improving smallholder's knowledge and adoption of sustainable land management (SLM) practices in agricultural production, the program employed two different modalities in the delivery of extension services. In the first, one farmer in each community was assigned to serve as a contact farmer for all other farmers in the community. The contact farmers received training in SLM practices as well as equipment. In the second modality, an additional female contact farmer was identified and she received the same training and equipment as the male contact farmer.

In addition to evaluating impact, Mogues et al. collected detailed data on costs incurred by IPs for the period of the project, which amounted to approximately \$3.7 million USD. The researchers found that the gendered treatment was more cost-effective than the program as a whole. The authors also simulated how the cost-effectiveness would change if the program were scaled up, and noted the

significant variation in the gains from scaling up, depending on the intervention modality. Comparing these two modalities in terms of their cost-effectiveness, as well as their impact, was particularly useful as the addition of a second female contact-farmer also brought with it higher costs than the basic intervention with only one male contact farmer per community.

In the context of LRG RCTs, cost-effectiveness analyses may be seen as an area of future research that has potential to provide additional value from rigorous evidence-based learning about land sector interventions, and contribute new knowledge to inform LRG policy and programmatic decision-making. Still, there are several limitations to conducting such studies, which may explain at least in part why they remain uncommon in LRG evaluations to date. In addition to the extra resources and time required to conduct CEA, in practice it is often very difficult for evaluation teams to collect cost data from IPs or make sense of information provided with respect to program effects. If the goal is to compare the effectiveness of two or more different approaches, it may also be challenging for IPs to develop alternative intervention approaches and roll them out simultaneously under the purview of a given program, unless the intervention alternatives entail fairly small differences. CEA typically may also rely on IPs agreeing to track costs for an independent third party in a particular way that is worked out with evaluation teams prior to program start. In practice this may be more feasible for smaller-scale interventions operating over fairly short time periods.

## **BLENDED METHODS**

RCTs typically provide highly credible evidence on the effects of development interventions. But, multiple complementary methods are often still needed to comprehensively understand how (the mechanisms), why, under what conditions, and for whom (what type of potential beneficiary) a given intervention leads to the desired set of outcomes and impacts. In other words, RCTs often still need to integrate qualitative data collection or other complementary research components, to understand the reasons why interventions work or not, for which types of beneficiaries, and under what broader context conditions.

RCTs can also be expensive, though often their cost is low relative to the cost of the interventions that are being evaluated. In addition, for the evaluation of many LRG interventions, the cost of an RCT evaluation design may in many cases be less than a difference-in-difference or next-best IE approach, because RCTs are often able to achieve similar or higher levels of statistical power with a smaller sample size, while the construction of a credible control group and the impact analysis and robustness checks for RCTs can also be more straightforward and less time consuming than for other IE approaches such as more commonly implemented difference-in-difference designs.

Critiques of RCTs have called for more intentional collaboration and integration of RCT approaches with other forms of quasi-experimental or non-experimental research under the same study umbrella, to help complement and strengthen the learning come out of the study. Doing so can also leverage strengths of each of those approaches, to better understand not just the impacts of a development intervention, but the mechanisms by which it worked, the reasons some beneficiaries may have benefited more than others, and how the implementation context and other factors may have influenced the outcomes (Bulte et al. 2020; Rodgers et al. 2020). These added benefits from blended approaches may also help to overcome skepticism over the RCT approach itself, make evaluation costs more

palatable, and enable decision-makers to utilize evaluation results and apply learning from the study more widely than might be possible from the RCT on its own.

## V. BEST PRACTICE GUIDANCE

### WHEN TO CONSIDER AN RCT OF AN LRG INTERVENTION

This report aimed to engage with a common challenge to conducting RCTs in the development sector, in which evaluation specialists may see a higher value-add to an RCT evaluation design over more complicated and potentially less rigorous and/or more costly quasi-experimental evaluation designs, while donors and development practitioners may be reluctant to explore this option even if an RCT design appears to be potentially feasible during an evaluation design phase. Reasons for this could include lower familiarity with RCT options, perceptions they are not feasible in development settings or have insurmountable ethical challenges, or other concerns about how to implement them and the nature of required collaborations. This may present a missed opportunity. In many situations the evidence-based learning and potential value-add from an RCT can be very compelling and potentially more credible than findings from other evaluation approaches, while perceived cost, logistical or potential ethical concerns may indeed be surmountable.

Development interventions often need to be complex, flexible and broad-based to achieve their multi-purpose aims within the contexts they are embedded in. At the same time, the types of interventions that are most amenable to RCTs tend to be those that are fairly narrow in focus and can utilize a standardized implementation approach throughout the intervention lifetime. Indeed, RCTs are often best suited to programs that are “*relatively simple... with clearly identified participants and non-participants, relatively short time horizons, and with little scope for the costs or benefits to spillover to the group of non-participants*” (Ravallion 2018). Such interventions may not always be reflective of the types of development programs for which learning is most needed, but this need not always be the case.

In keeping with decision factors for conducting RCTs in the development space more generally, RCTs of LRG interventions may be especially appropriate to consider when the following conditions are met:<sup>37</sup>

- There is likely to be a demonstrable value-add to learning beyond what could be obtained from other impact evaluation alternatives, such as from quasi-experimental approaches. For example, will the knowledge gained through the additional rigor of the RCT be likely to substantially advance current understanding of the presence and magnitude of impacts for a key type of intervention?;
- The intervention being considered for evaluation is well-defined and can be implemented with relatively low variability. In addition, IPs will not face overly burdensome challenges to maintain fairly standardized implementation of the intervention across different treatment units, and there are no inherent reasons to expect the intervention to vary considerably from place to place;
- Randomized assignment of treatment is possible logistically and politically, and is not likely to cause overly burdensome challenges to program implementation;
- Randomized assignment of treatment has a low probability of causing potential harms to beneficiaries or control group study participants;

---

<sup>37</sup> Dalziel, 2017.

- Randomized allocation of the intervention is not likely to reduce the treatment effects on participants due to variation in their ability or interest to benefit from the treatment;
- The planning for the intervention itself is still at an early enough stage, such that an evaluation team has sufficient time to scope out potential RCT approaches, discuss options collaboratively with IPs and USAID, finalize the design and any schedule or implementation requirements for the intervention and collect the evaluation baseline data before the planned start of the intervention. In practice, this typically means that the evaluation scoping work should start prior to or concurrent with the program award to IPs.

While RCTs may be a suitable option for many types of LRG interventions, it is important for evaluation teams and donors to consider the potential learning value-add of an RCT over alternative evaluation methods, and not simply assume that an RCT should always be done as the more superior option. Other questions that evaluation teams and donors should ask while considering RCT options include:

- *What implementation and other assumptions are associated with an RCT approach for a given intervention and evaluation design context?*
- *What is the specific treatment(s) or intervention(s) that could be tested through an RCT evaluation? (Particularly important to clarify when the program under evaluation has multiple components)*
- *What are the possibilities for conducting a more targeted RCT of one or more of those components within a broader program, or focusing on other elements of the intervention or how it is delivered that could help fill specific knowledge gaps about intervention effectiveness, causal mechanisms for how the intervention works or impacts for different sub-populations of interest?*
- *What is the potential added-value of the RCT from an evidence-based learning perspective, within the context of the current relevant knowledge and evidence base?*
- *How do the potential learning contributions mesh with USAID's learning priorities for the evaluation, and the potential contributions to the broader evidence base about the intervention or its theory of change?*

## **BEST PRACTICES FOR IMPLEMENTING RCTS OF LRG INTERVENTIONS**

In reviewing experiences from existing LRG RCTs, it is apparent that many of the challenges evaluation teams faced in implementing the RCT are no different from challenges that evaluation teams face in implementing other types of impact evaluations as well. Key issues that RCT approaches must additionally grapple with include gaining buy-in from IPs and other stakeholders, and working through the feasibility, potential challenges and mitigation options associated with selecting the unit at which the intervention will be assigned, and determining how the randomization will be accomplished. Below, we summarize key best practices that may help to strengthen the potential to design and carry out an RCT within the LRG space.

### **FOR USAID AND RELATED DEVELOPMENT AGENCIES**

USAID and other development agencies that are interested in funding LRG RCTs can help smooth the process when they:

- Develop clear learning priorities and articulate desired evidence needs for an independent evaluation of an intervention, either in conjunction with an evaluation team or prior to engaging the team. This can include visioning on how rigorous evidence-based learning from an evaluation might help to inform future programming decisions;
- Facilitate frequent and open communication and information-sharing between evaluation teams and IPs for the intervention under evaluation;

- Begin facilitating evaluation scoping and design work early in intervention planning.

### FOR EVALUATION TEAMS

Evaluation teams may be able to help strengthen the likelihood of an RCT design being accepted, and ensure smoother implementation and ultimate learning value when they:

- Can obtain a detailed understanding of the implementation context and intended program design from earlier in intervention planning;
- Conduct frequent outreach and clear communication with donors and IPs about potential design options and seek buy-in from early stages;
- Work to achieve a common vision and set of goals for the RCT across the various partners, incorporate learning interests from IPs where possible, and help to build a common understanding of the evaluation design and objectives;
- Aim to avoid requiring IPs to operate in contexts that are less familiar to them, or to implement the intended program to those who typically would not be prioritized for the planned intervention;
- To the extent possible, ensure that RCT designs employ methods to also shed light on why programs work or not and pathways to impact, rather than just focusing on obtaining a statistical measure of the intervention's effects. Often this entails integrating qualitative and non-experimental data collection and research methods into the study, and ensuring that these components draw on diverse expertise across disciplines and appropriately complement data from other aspects of the RCT, rather than being conducted in isolation;<sup>38</sup>
- Employ RCT approaches and sample designs that anticipate the possibility for spillover effects, partial treatment, or other threats to the evaluation validity, where possible;
- Consider ways to explicitly address heterogeneous treatment implementation,<sup>39</sup> and ensure a sufficiently powered design to examine heterogeneous treatment effects on different sub-populations of interest among the beneficiary population (for LRG interventions, this could include, for example, female-headed households; women within households; poorer or otherwise more vulnerable households; members of other particularly marginalized groups);
- Anticipate that obtaining buy-in and adherence by IPs and government or other stakeholders will likely require an iterative process, much advance planning, and working collaboratively with all partners to explain the process, weigh trade-offs and ensure feasibility;
- Recognize that RCT approaches do often impose some additional constraints on IPs with respect to how they implement a given intervention, in order to maintain the integrity of the RCT. Aim to work with IPs to discuss potential implementation permutations that IPs might envision, and explore possibilities to work with this or build this explicitly into the RCT design through multiple treatment arms;
- When possible, use active time on the ground with the IP team in-country, for example through scoping visits, data collection preparations and data collection itself, to firm up relationships and continue to strengthen collaboration throughout the evaluation;
- Maintain proactive and frequent communication with IPs regarding implementation progress once the intervention is underway;
- Involve IPs in all stages of the design and implementation of the RCT, but seek to minimize the additional burden on IPs to the extent possible.

---

<sup>38</sup> Menon et al, 2020.

<sup>39</sup> Carter and Barrett 2020.

## **FOR IMPLEMENTING PARTERS**

RCTs of LRG interventions are likely to proceed more smoothly and yield the intended evidence-based learning when IPs:

- Seek to understand the learning benefits of an RCT evaluation of the intervention and what the approach will entail;
- Consider how a rigorous evidence-based approach to learning from evaluation activities could be leveraged to provide additional learning for implementation or to help make course corrections as implementation proceeds;
- Work collaboratively with the evaluation team to communicate, share information, and help ensure coordination between the program implementation and the RCT design and implementation;
- Use knowledge of the RCT to help work with and keep other program stakeholders, such as local government authorities, informed about the study;
- Are open to and aim to follow the agreed RCT protocol with respect to program implementation. If potential challenges come up, reach out to the evaluation team early;
- Are proactive about reaching out to the evaluation team to discuss options when potential complications, delays, or other changes to planned implementation arise, or otherwise unexpected issues occur over the course of implementation that might be relevant for the RCT.

## LITERATURE CITED

- Ali, D. A., Collin, M., Deininger, K., Dercon, S., Sandefur, J., & Zeitlin, A. 2014. The price of empowerment: Experimental evidence on land titling in Tanzania. *Center for Global Development Working Paper*, (369).
- Ali, D. A., Collin, M., Deininger, K., Dercon, S., Sandefur, J., & Zeitlin, A. 2016. Small price incentives increase women's access to land titles in Tanzania. *Journal of Development Economics*, 123, 107-122.
- Angelucci, M.; Di Maro, Vincenzo. 2015. Program Evaluation and Spillover Effects. Policy Research Working Paper; No. 7243. World Bank, Washington, DC. © World Bank.  
<https://openknowledge.worldbank.org/handle/10986/21857> License: CC BY 3.0 IGO.
- Arnot, C. D., Luckert, M. K., & Boxall, P. C. 2011. What is tenure security? Conceptual implications for empirical analysis. *Land Economics*, 87(2), 297-311.
- Banerjee, A. V., Duflo, E., & Kremer, M. 2016. The influence of randomized controlled trials on development economics research and on development policy. *The State of Economics, The State of the World*, 482-488.
- Baylis, K. and A. Ham, 2015. How important is spatial correlation in randomized controlled trials? (No. 330-2016-13548). Available at: <https://ageconsearch.umn.edu/record/205586/>
- Baylis, K., Honey-Rosés, J., Börner, J., Corbera, E., Ezzine-de-Blas, D., Ferraro, P.J. et al. 2016. Mainstreaming impact evaluation in nature conservation. *Conservation Letters*, 9, 58–64.
- Bédécarrats, F., Guérin, I., & Roubaud, F. 2019. All that glitters is not gold. The political economy of randomized evaluations in development. *Development and Change*, 50(3), 735-762.
- Bennett, N. J., & Satterfield, T. 2018. Environmental governance: A practical framework to guide design, evaluation, and analysis. *Conservation Letters*, 11(6), e12600.
- Bulte, E., Di Falco, S., & Lensink, R. 2020. Randomized interventions and “real” treatment effects: A cautionary tale and an example. *World Development*, 127, 104790.
- Cameron, D. B., Mishra, A., & Brown, A. N. 2016. The growth of impact evaluation for international development: how much have we learned?. *Journal of Development Effectiveness*, 8(1), 1-21.
- Collin, M. E. 2017. Peer effects in the demand for property rights: experimental evidence from urban Tanzania. *World Bank Policy Research Working Paper*, (8163).
- Dalziel, M. 2018. Why are there (almost) no randomised controlled trial-based evaluations of business support programmes?. *Palgrave Communications*, 4(1), 1-9.
- de Souza Leão, L., & Eyal, G. 2019. The rise of randomized controlled trials (RCTs) in international development in historical perspective. *Theory and Society*, 48(3), 383-418.
- Deaton, A., & Cartwright, N. 2018. Understanding and misunderstanding randomized controlled trials. *Social Science & Medicine*, 210, 2-21.
- Deininger, K. W. 2003. *Land policies for growth and poverty reduction*. World Bank Publications.
- Donovan, K. P. 2018. The rise of the randomistas: on the experimental turn in international aid. *Economy and Society*, 47(1), 27-58.

- Fenske, J. 2011. Land tenure and investment incentives: Evidence from West Africa. *Journal of Development Economics*, 95(2), 137-156.
- Gertler, P. J., Martinez, S., Premand, P., Rawlings, L. B., & Vermeersch, C. M. 2016. *Impact evaluation in practice*. World Bank Publications.
- Ghebru, H., & Lambrecht, I. 2017. Drivers of perceived land tenure (in) security: Empirical evidence from Ghana. *Land Use Policy*, 66, 293-303.
- GIL, 2019. Impact Evaluation of Access to Land Project in Benin. World Bank Gender Innovation Lab.
- Glennerster, R. & Takavarasha, K. 2013. *Running Randomized Evaluations: a Practical Guide*. Princeton University Press, Princeton, USA.
- Goldstein, M., Houngbedji, K., Kondylis, F., O'Sullivan, M., & Selod, H. 2015. Formalizing Rural Land Rights in West Africa.
- Goldstein, M., Houngbedji, K., Kondylis, F., O'Sullivan, M., & Selod, H. 2018. Formalization without certification? Experimental evidence on property rights and investment. *Journal of Development Economics*, 132, 57-74.
- Graham, J., Amos, B., & Plumptre, T. W. 2003. Principles for good governance in the 21st century. Ottawa: Institute on Governance, Governance Principles for Protected Areas.
- Holden, S. T., & Ghebru, H. 2016. Land tenure reforms, tenure security and food security in poor agrarian economies: Causal linkages and research gaps. *Global Food Security*, 10, 21-28.
- Higgins, D., Balint, T., Liversage, H., and P. Winters. 2018. Investigating the impacts of increased rural land tenure security: a systematic review of the evidence. *Journal of Rural Studies*. 61:34–62.
- Hoffman, N. 2020. Involuntary experiments in former colonies: The case for a moratorium, *World Development*, 127, 104805.
- Huntington, H., Starosta, A. B. Ewing, and N. Walter. 2018. Tenure and Global Climate Change Evaluation Report. USAID. Washington, DC.
- Huntington, H., & Shenoy, A. 2021. Does insecure land tenure deter investment? Evidence from a randomized controlled trial. *Journal of Development Economics*, 150, 102632.
- IPA, 2016. Millennium Challenge Account of Mongolia (MCA-M) Peri-Urban Land Leasing Activity. Interim Report for Phase II Areas. MCC and IPA.
- Knight, R., J. Adoko, R. Auma, A. Kaba, A. Salomao, S. Siakor, and I. Tankar. 2012. Protecting Community Lands and Resources: Evidence from Liberia, Mozambique and Uganda. Namati and International Development Law Organization.
- Knight, R., J. Adoko, and R. Auma. 2013a. Protecting Community Lands and Resources: Evidence from Uganda. Namati, LEMU and International Development Law Organization.
- Knight, R., S. Siakor, and A. Kaba. 2013b. Protecting Community Lands and Resources: Evidence from Liberia. Namati, SDI and International Development Law Organization.
- Knight, R., A. Salomao, and I. Tankar. 2014. Protecting Community Lands and Resources: Evidence from Mozambique. International Development Law Organization, Centro Terra Viva and Namati.
- Lawry, S., Samii, C., Hall, R., Leopold, A., Hornby, D., & Mtero, F. 2017. The impact of land property rights interventions on investment and agricultural productivity in developing countries: a systematic review. *Journal of Development Effectiveness*, 9(1), 61-81.

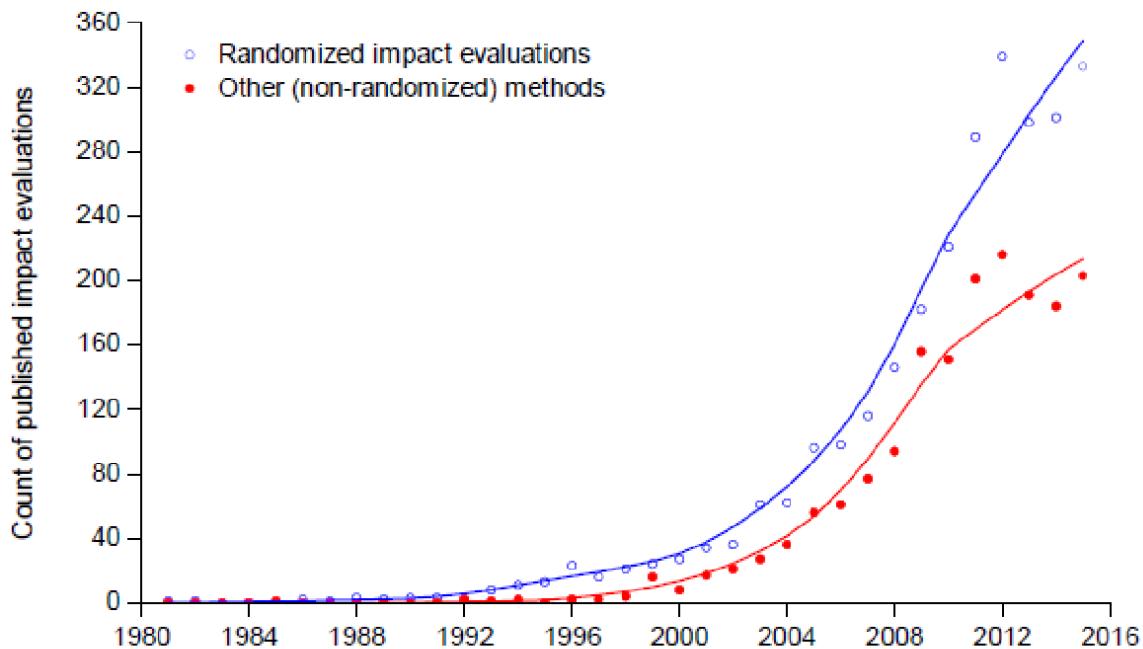
- Lisher, J. W. 2019. Guidelines for impact evaluation of land tenure and governance interventions. Global Land Tenure Network, Nairobi.
- Lockwood, M. 2010. Good governance for terrestrial protected areas: A framework, principles and performance outcomes. *Journal of environmental management*, 91(3), 754-766.
- Lockwood, M., Davidson, J., Curtis, A., Stratford, E., & Griffith, R. 2010. Governance principles for natural resource management. *Society and natural resources*, 23(10), 986-1001.
- Meinzen-Dick, R. S., Brown, L. R., Feldstein, H. S., & Quisumbing, A. R. 1997. Gender, property rights, and natural resources. *World development*, 25(8), 1303-1315.
- Meinzen-Dick, R., & Mwangi, E. 2009. Cutting the web of interests: Pitfalls of formalizing property rights. *Land use policy*, 26(1), 36-43.
- Meinzen-Dick, R., Quisumbing, A., Doss, C., & Theis, S. 2019. Women's land rights as a pathway to poverty reduction: Framework and review of available evidence. *Agricultural Systems*, 172, 72-82.
- Menon, P., M. T. Ruel, P. H. Nguyen, S. S. Kim, K. Lapping, E. A. Frongillo, S. Alayon. 2020. Lessons from using cluster-randomized evaluations to build evidence on large-scale nutrition behavior change interventions, *World Development*, 127, 104816.
- Mogues T, Mueller V, Kondylis F. 2019. Cost-effectiveness of community-based gendered advisory services to farmers: Analysis in Mozambique and Tanzania. *PLoS ONE* 14(3): e0211448. <https://doi.org/10.1371/journal.pone.0211448>
- Mueller, V., Billings, L., Mogues, T., Peterman, A., & Wineman, A. 2015. Filling the legal void? Experimental evidence from a community-based legal aid program for gender-equal land rights in Tanzania.
- North, D. 1990. Institutions, Institutional Change and Economic performance. New York: Cambridge University Press.
- Persha, L. and J. Patterson-Stein. 2021. Endline Evaluation Report: Impact Evaluation of the Feed the Future Tanzania Land Tenure Assistance Activity. E3 Analytics and Evaluation Project. USAID.
- Persha, L. and J. Patterson-Stein. 2018. Phase II Baseline and Phase I Midline Report: Impact Evaluation of the Feed the Future Tanzania Land Tenure Assistance Activity. United States Agency for International Development. Washington, DC.
- Place, F. 2009. Land tenure and agricultural productivity in Africa: a comparative analysis of the economics literature and recent policy strategies and reforms. *World development*, 37(8), 1326-1336.
- Rasolofoson, R.A., Ferraro, P.J., Jenkins, C.N. & Jones, J.P.G. 2015. Effectiveness of community forest management at reducing deforestation in Madagascar. *Biological Conservation*, 184, 271–277.
- Ravallion, M. 2018. Should the Randomistas (Continue to) Rule? CDG Working Paper 492.
- Ravallion, M. 2020. Highly prized experiments. *World Development*, 127, 104824.
- Reyes-García, Victoria, et al. 2012. Does participatory mapping increase conflicts? A randomized evaluation in the Bolivian Amazon. *Applied Geography* 34: 650-658.
- van der Meulen Rodgers, Y., Bebbington, A., Boone, C., Dell'Angelo, J., Platteau, J. P., & Agrawal, A. 2020. Experimental approaches in development and poverty alleviation. *World Development*, 127, 104807.

- Sabet, S. M., & Brown, A. N. 2018. Is impact evaluation still on the rise? The new trends in 2010–2015. *Journal of Development Effectiveness*, 10(3), 291-304.
- Sandefur, J., & Siddiqi, B. 2013. Delivering justice to the poor: theory and experimental evidence from Liberia. In *World Bank Workshop on African Political Economy, Washington, DC, May* (Vol. 20).
- Singer, P., Baker, A., & Haushofer, J. 2019. Are randomized poverty alleviation experiments ethicals. *Project Syndicate*.
- Snistveit, B., J. Stevenson, P. Fenton Villar, J. Eyers, C. Harvey, S. Panfil, J. Puri, and M. C. McKinnon. 2016. Land use change and forestry programmes: What we know and don't know about the effects on greenhouse gas emissions and food security.
- Stevens, C., Panfil, Y., Linkow, B., Hagopian, A., Mellon, C., Heidenrich, T., Kulkarni, N., Bouvier, I., Brooks, S., Lowery, S., & Green, J. 2020. Land and Development: A Learning Agenda for Land and Resource Governance at USAID. Washington, DC.
- Tseng, Tzu-Wei Joy, Brian E. Robinson, Marc F. Bellemare, Ariel BenYishay, Allen Blackman, Timothy Boucher, Malcolm Childress et al. 2021. Influence of land tenure interventions on human well-being and environmental outcomes. *Nature Sustainability* 4, no. 3: 242-251.
- USAID. 2011. USAID Evaluation policy: Evaluation, learning from experience. USAID. Washington, DC.
- White, H. 2010. A contribution to current debates in impact evaluation. *Evaluation*, 16(2), 153-164.
- White, H. 2013. An introduction to the use of randomised control trials to evaluate development interventions. *Journal of Development Effectiveness*, 5, 30–49.
- Winters, P., Salazar, L., & Maffioli, A. 2010. Designing impact evaluations for agricultural projects. *Inter-American Development Bank, Washington, DC*, 14, 2012.
- Wren-Lewis, L., Becerra-Valbuena, L., & Houngbedji, K. 2020. Formalizing land rights can reduce forest loss: Experimental evidence from Benin. *Science Advances*, 6(26), eabb6914.

## ANNEX I. INCREASE IN DEVELOPMENT RCTs OVER TIME

Several authors have highlighted an increase in recent decades in the use of impact evaluations as a tool for learning about development program impacts, together with a rise in RCT evaluations specifically. The figure below, from Ravallion 2018,<sup>40</sup> provides an illustration of this by showing an increase in published evaluations that use an RCT approach, drawing on studies catalogued in the International Initiative for Impact Evaluation's (3ie) evaluation database.

**Figure 1. Annual counts of published impact evaluations for developing countries**



Note: Fitted lines are nearest neighbor smoothed scatter plots. See footnote 4 in the main text on likely under-counting of non-randomized IE<sup>s</sup> in earlier years. Source: International Initiative for Impact Evaluation.

(Figure Source: Ravallion 2018)

<sup>40</sup> Ravallion, M. 2018. Should the Randomistas (Continue to) Rule? CDG Working Paper 492.

## ANNEX II. RCT DESIGN CONSIDERATIONS AND DECISION-TREE

An RCT design can sometimes be more feasible than you might think on the surface. Select the right ingredients based on learning objectives and the realities of program implementation, to customize an RCT suitable for the intended purposes. See different strategies below to adapt RCTs to evaluate land programming. These can be mixed and matched as the circumstances dictate.

### CLUSTER RANDOMIZATION



**What is it?** Instead of randomly selecting individuals or households for the intervention, randomly select whole clusters (e.g. villages, districts) to receive intervention.

**When should you use it?** The intervention must be applied to whole clusters at once; and/or there is high risk of neighboring households sharing access to the intervention if it is withheld.

**Any tips for use?** Cluster randomization requires a larger sample to reach conclusions with the same degree of certainty as individual-level randomization. Plan for at least 30-50 clusters, and many households sampled in each.

#### Land sector example:

A 2016-2021 RCT of the USAID/Tanzania LTA program worked with district authorities to determine which villages would receive the program. Four years later, with positive results from the program, LTA expanded services to the remaining eligible villages.



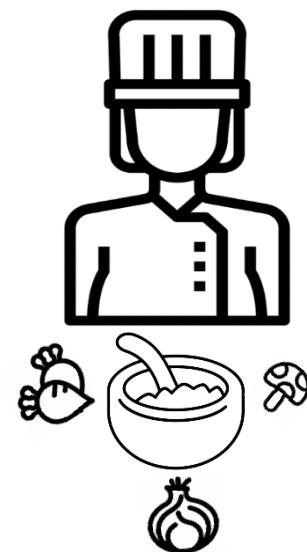
### PHASED/PIPELINE DESIGN



**What is it?** Instead of randomizing the set of participants who will receive an intervention, stagger timing of the intervention for all intended participants.

**When should you use it?** You would like all potential participants to receive the intervention; there is insufficient budget to treat the entire population at once.

**Any tips for use?** Factor timing of intended outcomes into length of the staggered roll-out. For example, if anticipated outcomes are expected in one year, wait at least one year until implementing the second wave of intervention.



### RANDOMIZE PROGRAM COMPONENTS



**What is it?** Instead of randomizing full intervention, randomize some component of intervention (e.g. different levels of subsidy, training plus technical assistance vs. training alone, etc.)

**When should you use it?** Efficacy of full intervention is known or obvious, but tradeoffs between possible components of program are unclear.

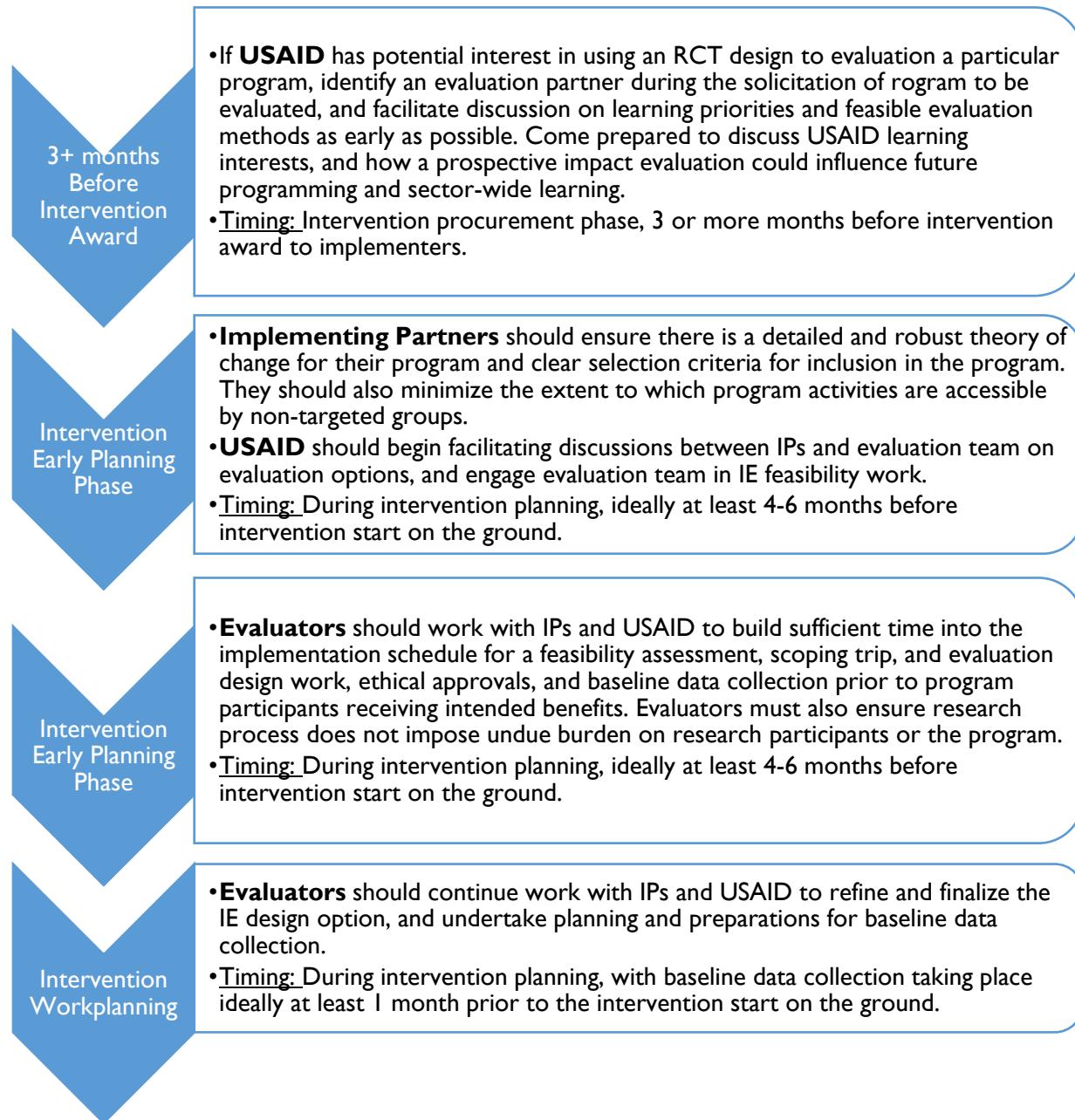
**Any tips for use?** Map the randomized component to a specific learning interest. Consider whether the standard program is a sufficient control, or whether a true control group (i.e. no intervention) is still needed.

#### Land sector example:

A 2019 RCT of a World Bank program in Uganda tested the stand-alone and complementary impacts of land titling and improved access to credit at the village and household levels in rural areas of Uganda. The evaluation further tested gender-specific effects of different strategies to promote land titling.



## KEY TIMELINE CONSIDERATIONS FOR DETERMINING RCT FEASIBILITY



## PRACTICAL TIPS FOR CONDUCTING AN RCT

Best practice ways to address concerns about RCTs include transparent communication with program implementers and related stakeholders about the reasons for randomization, potential benefits, and randomization options that might be available. It is also important for evaluation teams to provide sufficient information and seek buy-in about the potential for an RCT design from early in the evaluation design process. There is a need to maintain close involvement of these stakeholders throughout the planning. USAID can play an essential role by helping to facilitate this communication from early in the evaluation design phase, and by clarifying for all USAID's main learning priorities out of the evaluation.

RCTs of LRG interventions may be appropriate to consider when the following conditions are met:

There is likely to be a demonstrable value-add to learning beyond what could be obtained from other impact evaluation alternatives, such as from quasi-experimental approaches. Will the knowledge gained through the additional rigor of the RCT likely to advance current understanding of impacts from the intervention?

The intervention being considered for evaluation is well-defined and can be implemented with relatively low variability. In addition, IPs will not face overly burdensome challenges to maintain fairly standardized implementation of the intervention across different treatment units, and there are no inherent reasons to expect the intervention to vary considerably from place to place.

Randomized assignment of treatment is possible logistically and politically, and unlikely to cause overly burdensome challenges to program implementation.

Randomized assignment of treatment has a low probability of causing potential harms to beneficiaries or control group participants.

Randomized allocation of the intervention is not likely to reduce the treatment effects on participants due to variation in their ability or interest to benefit from the treatment.

The planning for the intervention itself is still at an early enough stage, such that an evaluation team has sufficient time to scope out potential RCT approaches, discuss options collaboratively with IPs and USAID, finalize the design and any schedule or implementation requirements for the intervention and collect the evaluation baseline data before the planned start of the intervention. In practice, this typically means that the evaluation scoping work should start prior to or concurrent with the program award to IPs.

Other questions that evaluation teams and donors should ask while considering RCT options include:

- *What implementation and other assumptions are associated with an RCT approach for a given intervention and evaluation design context?*
- *What is the specific treatment(s) or intervention(s) that could be tested through an RCT evaluation? (Particularly important to clarify when the program under evaluation has multiple components)*
- *What are the possibilities for conducting a more targeted RCT of one or more of those components within a broader program, or focusing on other elements of the intervention or how it is delivered that could help fill specific knowledge gaps about intervention effectiveness, causal mechanisms for how the intervention works or affects different sub- populations of interest?*
- *What is the potential added-value of the RCT from an evidence-based learning perspective, within the context of the current relevant knowledge and evidence base?*
- *How do the potential learning contributions mesh with USAID's learning priorities for the evaluation, and the potential contributions to the broader evidence base about the intervention or its theory of change?*

## DECISION TREE FOR RCT ALTERNATIVES

If no variation of an RCT is possible, other evaluation designs are still possible which can also yield credible evidence on the impacts of USAID land programming. The decision tree below elaborates how to select which evaluation design is most suitable.

