Evaluating the impact of Land Administration Programs on agricultural productivity and rural development

Jeremie Gignoux
Karen Macours
Liam Wren-Lewis

Inter-American Development Bank

Office of Strategic Planning and Development Effectiveness (SPD) and Environment, Rural Development and Disaster Risk Management Division (RND)

TECHNICAL NOTE

No. IDB-TN-506

January 2013

Evaluating the impact of Land Administration Programs on agricultural productivity and rural development

Jeremie Gignoux
Karen Macours
Liam Wren-Lewis

Inter-American Development Bank
2013

Cataloging-in-Publication data provided by the
Inter-American Development Bank
Felipe Herrera Library

Gignoux, Jeremie.
Evaluating the impact of land administration programs on agricultural productivity and rural development / Jeremie Gignoux, Karen Macours, Liam Wren-Lewis.
p. cm. (IDB Technical Note; 506)
Includes bibliographical references.
1. Land use. 2. Agricultural productivity. 3. Rural development. 4. Agricultural development projects—Evaluation. I. Macours, Karen. II. Wren-Lewis, Liam. III. Inter-American Development Bank. Office of Strategic Planning and Development Effectiveness. IV. Title. V. Series.
IDB-TN-506

http://www.iadb.org

The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the Inter-American Development Bank, its Board of Directors, or the countries they represent.

The unauthorized commercial use of Bank documents is prohibited and may be punishable under the Bank's policies and/or applicable laws.

Copyright © 2013 Inter-American Development Bank. All rights reserved; may be freely reproduced for any non-commercial purpose.

Jeremie Gignoux, Paris School of Economics (PSE) and French National Institute for Agricultural Research (INRA), 48 boulevard Jourdan, 75014 Paris, gignoux@pse.ens.fr. Karen Macours, Paris School of Economics (PSE) and French National Institute for Agricultural Research (INRA), 48 boulevard Jourdan, 75014 Paris, karen.macours@parisschoolofeconomics.eu. Liam Wren-Lewis, Paris School of Economics (PSE) and French National Institute for Agricultural Research (INRA), 48 boulevard Jourdan, 75014 Paris, liamwrenlewis@gmail.com.

Evaluating the impact of Land Administration Programs on agricultural productivity and rural development

Jeremie Gignoux, Karen Macours and Liam Wren-Lewis[1]

Abstract: Investment in land administration projects is often considered key for agricultural productivity and rural development. The evidence on such interventions is however remarkably mixed. This paper discusses a number of challenges and derives related guidelines for the impact evaluation of land administration programs. We focus on four types of challenges: 1) a conceptual challenge related to the need to unbundle property rights and to establish the plausible causal chain for land administration interventions; 2) the existence of other binding constraints on productivity, implying the need to consider the complementarity between property rights and other productive interventions; 3) methodological challenges related to the causal identification of the impacts of such interventions; 4) practical and operational challenges for good impact evaluation on land administration, which are often operationally complex and politically sensitive. The paper is specifically written with a focus on land administration projects in Latin America, and draws from experiences and lessons learned from impact evaluations designed for the IDB, and other donors-funded land administration projects in the region.

JEL codes: D23, H43, O13, Q15

Keywords: Land administration programs, property rights, agricultural productivity, rural development, impact evaluation methods.

[1] Gignoux, Macours and Wren Lewis are at the Paris School of Economics (PSE) and French National Institute for Agricultural Research (INRA), 48 boulevard Jourdan, 75014 Paris, Contacts: gignoux@pse.ens.fr, karen.macours@parisschoolofeconomics.eu, liamwrenlewis@gmail.com.

This study was prepared for and funded by the Environment, Rural Development and Disaster Risk Management Division and the Office of Strategic Planning and Development Effectiveness, Inter-American Development Bank (IDB). It benefited from an impact evaluation project prepared in collaboration with Gilles Damais, Michele Lemay, and Marion LePommelec at the Haiti IDB country office. Marlon Seror provided excellent research assistance. We are grateful to Leonardo Corral, Lina Salazar, Jolyne Sandjak, Paul Winters and to participants at a workshop at the IDB for helpful comments on earlier drafts. We are solely responsible for any remaining errors.

1. Introduction

Insecurity of property rights is often argued to be an important impediment for agricultural productivity, and indeed more broadly for economic growth and prosperity. Land administration programs are designed to address such property rights insecurity, aiming at strengthening the rights of existing owners through clarification and formalization of individual rights, legislative changes, and/or improvements in conflict resolution mechanisms. If property rights security indeed leads to productivity gains, such interventions seem to hold a lot of promise. And as recent spikes in food prices have brought renewed attention to interventions that can increase agricultural productivity and food security, land administration programs might seem an attractive avenue for further investment. Yet while the Inter-American Development Bank (IDB) and other donors have invested in titling and other land administration programs in Latin America for a relatively long time, rigorous quantitative evidence on the impact of such interventions is rare, in particular for rural areas.

Even more remarkably, the existing evidence is very mixed. The standard theoretical argument is that property rights can affect agricultural productivity through investment, credit and land allocation (Feder and Feeny, 1991; Besley 1995). But empirical evidence on impacts of land administration on investment and land allocation is inconclusive, and the existing evidence for credit, if anything, mostly suggests no impact. As a consequence, empirically it is far from clear whether, how, and to what extent such programs can contribute to improving agricultural productivity.

This paper discusses a number of challenges related to the evaluation of land administration programs that, we believe, can in part explain the mixed evidence, and the confusing implications that can be derived from them. We do so in order to derive guidelines regarding the design of impact evaluations for new land administration programs. We focus on four types of challenges: 1) a conceptual challenge related to the need to unbundle property rights and to establish the plausible causal chain for a land administration interventions; 2) the existence of other binding constraints on productivity, implying the need to consider the complementarity between property rights and other productive interventions; 3) methodological challenges related to the causal identification of the impacts of such interventions; 4) practical and operational challenges for good impact evaluation on land administration, which are often operationally complex and politically very sensitive.

By discussing how these challenges can be addressed this paper aims at providing useful guidance for the design of impact evaluations for land administration projects. While detailed technical guidance for impact evaluations can be found in many other places, this paper aims to focus on some key aspects that are particularly pertinent for land administration projects. As such it aims to complement the closely related paper on impact evaluation for land property rights reform by Conning and Debb (2007). In particular, we emphasize the gains from integrated intervention and evaluation designs that unpack property rights in order to assess the incrementality of interventions. We further emphasize the complementarity with other types of interventions, for example smart subsidies to promote technology adoption. The paper also differs in its focus on outcomes and considers a wide range of intermediate and final outcomes.

The paper is specifically written with a focus on land administration projects in Latin America, where the starting position is typically one of existing land administration systems based on the Napoleonic code, with existing individual rights, that may be imperfect and insecure, due to e.g. overlapping and or outdated titles, and family co-ownership. While this has generally been ignored in the literature, it is unclear to what extent lessons based on evidence from elsewhere are relevant for Latin America, as the reasons for insecurity in other regions and hence the starting positions for land administration interventions are often quite different. For example, the underlying causes of land rights insecurity in Sub-Saharan Africa, where groups often have strong pre-existing rights, and where there might be a complete absence of a systematic land administration system, can be very different than in Latin America. While many of the challenges we discuss in this paper, such as the need to unpack property rights as well as the methodological challenges can also be relevant for land administration projects in other parts of the world, the types of interventions we discuss, are those with particular relevance for Latin America. We draw from experiences and lessons learned from impact evaluations designed for the IDB, and other donors-funded land administration projects in the region. In discussing methodological challenges, we also bring in studies from other regions, in case they provide useful examples.

The paper is organized as follows: section 2 discusses conceptual challenges for land administration impact evaluations focusing on trying to understand better what it is that such evaluations are trying to evaluate, and which outcomes to measure. We discuss how establishing a credible causal chain can provide lessons that go beyond the particular intervention that is

being evaluated. We further discuss how complementarity with other interventions can be crucial to understanding the full impact of land administration interventions. In section 3, we then turn to the methodological challenges, discussing several methods to investigate the impact of property rights security. We discuss both experimental and non-experimental methods, highlighting the challenge of addressing selection bias and trying to establish that the observed correlations can be interpreted as causal. We discuss various papers and indicate which methods have been or could be used to assess and to reduce possible biases. We also discuss data requirements for rigorous evaluations. Section 4 discusses the case study of an evaluation design in Haiti, as it allows illustrating several of the previous points. Finally in section 5, we conclude with summarizing the key practical insights and take home messages.

2. Conceptual challenges

2.1 The causal chain

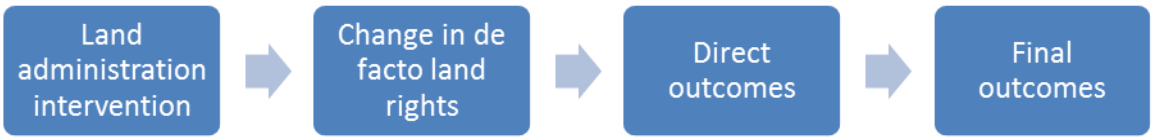
There are different types of land administration interventions that one may be interested in evaluating. IDB land administration projects typically contain a subset of the following activities:

(i) facilitation of broad-based land policy;
(ii) institutional strengthening (including decentralization) of the cadastral and registration agencies;
(iii) systematic regularization (saneamiento) of all land parcels in a given area;
(iv) introduction of digital information technology (IT) for managing cadastral and registry information;
(v) cadastral survey and mapping, happening typically at a disaggregate level;
(vi) property registration and land titling.

Most projects will combine a number of these interventions.

In assessing the impact of these projects on the outcomes of interest, there are several `links in the chain' that should be evaluated as part of any study on land administration interventions. These steps are outlined in Figure 1.

Figure 1: The chain from intervention to outcomes



An impact evaluation would aim to answer the following questions corresponding to the three links in the chain:

(i) What are the changes in land rights? How has the intervention changed these rights?

(ii) What are the direct intermediate outcomes? How have the changes in land rights led to changes in these intermediate outcomes?

(iii)What are the final outcomes?

Hence trying to establish these links between any particular intervention (or components thereof) and the theoretical arguments on expected impacts should be a first step in designing evaluations. This exercise should be done considering potential hypotheses of impacts at various levels of aggregation, including household-level impacts and municipal-level impacts (see further below). In this way, the evaluations can ultimately be more useful, as they can allow the drawing of lessons beyond the narrow experience of a particular project. This holds for each of the specific links in the causal chain.[2]

For instance, with regard to the first link, the previous theoretical literature has typically modeled `security enhancing reforms' as a decrease in the expected probability of expropriation. However, it has generally not modeled exactly how the government (or other actor) goes about generating this decrease. This absence seems notable given that experience suggests a simple change in the law is neither necessary nor sufficient to change the expected probability of expropriation. Perhaps partly resulting from this theoretical absence, there appears to be a lack of evaluations that identify exactly how a reform impacts perceived security. This is true even if it is through this `security' channel that most of the positive effects of titling are assumed to occur. Understanding the mechanisms through which reforms and land administration programs increase security is likely to be useful in designing and estimating the effect of future programs.

[2] This is directly linked to the preparation of the Results Matrix, as the ex-ante consideration of the theory of change and the listing of these hypotheses will help to establish the causal logic of the program with the anticipated results at each stage.

Identifying the channels through which a particular land administration project operates is also important to derive hypotheses related to the outcomes that can be expected to change (the second link). And almost equally important, it should also allow deriving hypotheses about which outcomes are unlikely to be affected. For instance, while higher security might well increase investments, its impact on access to credit is ambivalent (see above). And impacts on credit are likely to be higher when a particular land administration intervention also affects transferability. Failure to make this type of distinction, which is common in the literature, makes it hard to draw more generalizable lessons from any particular evaluation.

Overall then, to maximize lessons from an evaluation it helps to focus on measuring the actual mechanisms leading to the changes in outcomes, in addition to quantifying the changes in outcomes themselves. When designing impact evaluations, project leaders should consider explicitly how the evaluation will enable us to answer each of the questions above regarding all three parts of the chain. Answering the `how' questions may require complementing the quantitative impact evaluation design and related data collection with more qualitative data. An additional approach that may also be useful is to exploit heterogeneity and complementarities with other interventions, which we will discuss later in this section.

2.2 Unpacking property rights: what are we evaluating?

Let us now consider theoretically the first step of the chain – that is, how our interventions of interest may change effective land rights. There are several channels through which interventions can change property rights. Property rights can be imperfect or incomplete in many different ways. Indeed, land rights are made up of a bundle of different rights, including the right to use land, the right to derive income from it, and the right to sell, and each of those rights can be imperfect. In order to understand how land administration interventions may have an impact, it is useful to distinguish how they might affect this bundle of land rights along three different dimensions: i) increased expected security; ii) individualization of land rights; and iii) facilitation of transfer. A land administration intervention can lead to:

1. *Increased expected security*

   This includes any part of an intervention that makes the land rights less likely to be expropriated or contested, or that reduces the perceived likelihood of such events. Here we use expropriation to mean any transfer without the owner's consent, with two typical forms

of such transfer being to squatters/tenants, or someone else at the behest of the government (local or central). Interventions may increase the security of all land rights (e.g. by reducing the potential of conflict), or just of certain plots that were previously contested (e.g. by issuing plot specific documents).

2. Individualization of land rights

This includes any aspect of an intervention that transfers rights from groups (e.g. families, communes, or the state) to individuals. This individualization may be an individualization of usage rights, income rights or transfer rights. In Latin America the latter two are typically most common, with individualization likely to take one of three forms:

i. Individualization of family-owned land, i.e. land in co-ownership after inheritance.[3]
ii. Individualization of government owned land that is already being used by an individual.
iii. Individualization of communal or collective land that is already being used by an individual.

3. Facilitation of consensual ownership transfer

This includes any aspect of an intervention that reduces the transaction costs involved for owners of land to sell their land or control its inheritance.

Some land administration interventions may only have impact through one of these channels. For instance, an increase in the capacity of the department responsible for land transfers may simply facilitate ownership transfer without increasing its security or individualizing any group held rights. However, in practice, land administration projects can often change the nature of the property rights in several ways simultaneously. For example, land titling could potentially operate through all three channels: security may be enhanced if titles increase the enforcement of existing individual or group rights; rights could become more individual if the previous *de facto* arrangement was to treat the rights as belonging to a group (e.g. the family); and transfers may be facilitated if these are allowed by law but were previously prevented due to uncertainty. Which channels a particular project or policy works through will be determined by the specific components of the intervention, but also to a large extent by the country context, including what

[3] Note that this is in fact land that was fully individualized in the past (and often may have an individual though outdated title on the name of the ancestor).

land rights were previously in effect. For instance, land titling can only facilitate transfers in a context where transfers are legal, and will only individualize rights when group rights were the previous *de facto* arrangement.

Table 1 below gives a potential mapping of the intervention types described above to the channels they are likely to work through. In each section, we describe part of the necessary conditions for a particular type of intervention to act through each of the three channels identified above. However, the channels that operate will be very dependent on the exact nature of the intervention and the context in which it operates. A key first step in designing an impact evaluation is therefore to identify the conditions under which each channel may (or may not) operate. This will help the impact evaluation be tailored to the potential mechanisms at work, and also help in identifying heterogeneity that could be useful to exploit in understanding the intervention's impact.

The theoretical literature has typically started from the point of one of these three channels, with little work analyzing how interventions work through these channels. In other words, economic theory has worked mostly on the question "what is the impact of a change in land rights?" and typically ignored the question "how do land administration interventions affect land rights?" This absence seems notable given that experience suggests a simple change in the law is neither necessary nor sufficient to change effective rights. An exception is Castañeda Dower and Pfutze (2012), who model how certification may enable a community to coordinate on enforcing a regime with different expropriation rules, and hence increase expected security.

Perhaps partly resulting from this theoretical absence, there are few evaluations that explore how a particular intervention changes land rights. Most empirical papers study interventions which are intended to increase expected security, but there are a number of ways in which a given intervention may do so. Possible mechanisms include:

(i) Providing information to land users on their existing rights

(ii) Reducing the cost and/or increasing the expected probability of success in invoking central government enforcement in the case of future conflict (i.e. the courts)

(iii) Coordinating local enforcement mechanisms

(iv) Reducing the expectation of future land reform and/or government expropriation

Table 1: Mapping of projects to possible channels

	Increased expected security	**Individualization of land rights**	**Facilitation of consensual ownership transfer**
Legal and policy changes	Possibly, for instance changes facilitating ownership registration or verification, or changes reducing expropriation risk.	Possibly for instance changes granting individual use rights on communal land.	Possibly, for instance granting of rights to rent and/or sell plots of land for which only use rights were previously held.
Institutional strengthening (including IT)	If previous institutional weakness led to lack of enforcement. Even then, effect is likely to be slow unless accompanied by information campaigns.	If previous weaknesses led (in some cases) to use of group rights, and strengthened institutions enforce individual rights.	If strengthened institution allows such transfers, and either (a) formal registration of transfers is easier, or (b) relative enforcement of land that's changed ownership is stronger
Systematic regularization	Yes	If `irregular' system involved use of group rights If transfer of regularized land is permitted and relatively cheap / easy	If transfer of regularized land is permitted
Introduction of digital information technology (IT) for managing cadastral and registry information	Yes, on longer term by keeping cadastral information updated	No	Yes, by facilitating updating of cadastral information reflecting transfer
Cadastral survey and mapping	Possibly, if mapping strengthens existing informal rights	No, unless through sub-division of land in family co-ownership	Possibly, if mapping strengthens existing informal rights
Land-titling	Yes	Only if previous system involved use of group rights	If transfer of titles is permitted and relatively cheap / easy
Registration	If formal enforcement mechanisms are stronger than informal ones	If formal enforcement mechanisms are stronger than informal ones, and informal mechanisms enforce group rights more than formal ones	If transfer of registration is permitted and relatively cheap / easy, and formal enforcement mechanisms are effective

Each of these effects of an intervention may increase expected security. Establishing which is at work is important when designing impact evaluations of land administration interventions, since it will provide information as to which aspects of the intervention are most effective, and give an idea as to the applicability of the intervention to other situations.

While the above discussion is general, in terms of methodological approaches for the impact evaluation, it is useful to distinguish between the various components or interventions any given land administration project might have. The interventions that occur at the national level (such as facilitation of national debates or changes in centralized processes) are typically hard to evaluate quantitatively, as their national scope prevents the establishment of a credible counterfactual. For those components, good qualitative process evaluations are best suited to shed light on their functioning and potential outcomes. However, for components that do not have to affect all land at once, such as cadastral mapping or titling, a quantitative evaluation strategy can be developed. And for some components that might in theory affect everybody (such as changes in land legislation), complementary information campaigns can sometimes be used to increase information about such changes for a subset of the population, allowing to measure their impact (see encouragement design below). Alternatively comparative case studies methods such as synthetic cohorts may provide evidence when some subgroup (e.g. a few regions) is affected only. For projects with different components, an identification of which of those components to evaluate quantitatively is hence a first step. In the rest of this paper, we focus on strategies for quantitative evaluations of land administration interventions that do not simultaneously take place nationwide.

2.3 Which outcomes?

The empirical literature has considered a wide variety of outcomes when evaluating impacts of land rights. Most of the focus has been on direct intermediate outcomes, hypothesized by the theoretical literature to be affected through one of the three main channels outlined above. We consider those first, prior to looking at final outcomes, where we distinguish between outcomes at the households-level and outcomes at a municipal or higher level of aggregation.

2.3.1 Direct (Intermediate) outcomes:

There are five potential direct outcomes of land administration interventions: Investment, credit, transfers of effective rights, time allocation/labor and migration, and conflict. Table 2 displays some of the key theoretical mechanisms through which changes to land rights can impact these outcomes. Appendix 1 and Table 1 in the appendix give an overview of the related empirical evidence.[4] We discuss each of these five impacts below.

(i) Investment

Perhaps the most frequently cited benefit of land administration interventions is that increased security will increase the expected time horizon of land-users and hence increase their investment. This may include the planting of trees, purchase of machinery or keeping land fallow. It may also include the choice of which crops to plant, since certain crops might 'invest' nutrients in the soil or provide seeds for future use. A slightly more subtle reason why investment may change is that certain investment activities may directly influence the probability of expropriation. For example, leaving land fallow may increase the expropriation probability, or the planting of trees may reduce the probability. Hence interventions that increase security may reduce the need for these security-enhancing actions (Sjaastad and Bromley, 1997; Goldstein and Udry, 2008). On the other hand, a potentially negative effect of increasing owners' security on investment is that this may reduce the investment incentives of tenants who were hoping for beneficial expropriation (Banerjee and Ghatak, 2004; Besley and Ghatak, 2009). Individualization may also increase investment by reducing moral hazard (Alchian and Demsetz, 1972) and the tragedy of the commons (Hardin, 1968). Finally, transfer facilitation may also increase the expected time horizon, and hence investment, if it enables land to be passed on to a designated heir or sold on a market. The intervention may change both short and long-term investments, and though the direct impact is likely to be mostly on agricultural investment, non-agricultural investment may also be affected. While generally perceived to be positive, the direction of the effect will be ambiguous, as it will depend on the overall investment portfolio of the household, and how the investment decision reacts to the expected changes in returns that the land administration intervention will bring.

[4] Table 2 illustrates that most of the available evidence relates to either intervention that clearly work through tenure security, or interventions that offer a combination of tenure security, individualization and transfer facilitation, but for which the empirical work does not allow to distinguish between these channels.

(ii) Credit

Following the work of de Soto (2000) and others, it is hypothesized that interventions that facilitate the transfer of land to financial institutions and subsequent land transactions will increase the ability of landowners to receive credit. Credit may also be increased through security enhancing actions, since this may increase the value of land to these financial institutions, as well as increasing the demand for credit by giving owners greater investment incentives.

(iii)Transfers of land rights

Land administration interventions may impact the frequency and nature of three types of land rights transfers: Sales, rentals and non-financial transfers. Many interventions aim to facilitate such transfers by reducing the involved transaction costs.

a. ***Sales***

Facilitation of rights transfer often focuses on improving the market for ownership rights. This is hoped to increase the transfer of land to owners who have a relative advantage, through exploiting economies of scale or a greater capacity for investment. On other hand, land may also be transferred to those looking to store value and those who are less risk-averse, which may not necessarily be welfare enhancing (Deininger and Feder, 2001). The frequency of sales may also increase if owners use land as a liquid asset to smooth consumption. Interventions that increase expected security may increase sales since they are likely to increase the security of potential purchasers more than that of existing owners.

b. ***Rentals***

Interventions that facilitate land transfers may directly affect land rental markets. Furthermore, increased ownership security may make owners less fearful of renting out their land, and hence reduce the expected transaction costs of rentals. This reduction in transaction costs may also lead to rental contracts of longer duration, with more diverse partners and under different contract types.

c. ***Non-financial transfers***

Though land administration interventions are not typically aimed at transfers outside of market processes, this may be a potential impact. In particular, land administration interventions may transfer effective rights within the household, thereby empowering

women or giving latter generations enhanced inheritance rights. An unintended consequence may also be the transfer of rights in cases of conflict, where rights may be gained by those that can use the intervention to their advantage.

(iv) Time allocation and migration

Greater security of ownership is expected to reduce the need to spend time on the land in question, and hence increase the time household members spend on other activities. This may include greater labor market participation and, in the extreme, migration away from the land in question. Transfer facilitation may also reduce the time spent by landowners on the land, since owners may be more able to sell or rent out their land rather than work on it themselves.

(v) Conflict

Greater security of ownership may reduce conflict over land, since the increased certainty should decrease the payoffs of fighting over land. Indeed, the process of providing greater security, e.g. through rights clarifications, might explicitly include efforts to resolve existing conflicts. Individualization may also reduce conflict amongst groups that previously jointly held rights to a piece of land, since the process clarifies the rights of individuals that may previously have been fought over. However, to the extent that stakeholders expect to see their claims recognized, the announcement of a clarification or individual titling intervention may spark latent conflicts in the short run. Facilitation of transfer, may have ambiguous effects even in the long run. On the one hand, a greater set of potential transfers may help to resolve conflicts in ways that were previously not possible. On the other hand, the greater possibility of transfer may increase the returns to conflict for non-owners, as well as allow transfers over which there is discord.

2.3.2 Household-level final outcomes

Whilst the above impacts are those that are most likely to result directly from land administration programs, they are not normally the ultimate objective of such interventions. Typically, it is hoped that positive impacts on the above direct variables will lead to improved final outcomes for farms and households. These outcomes principally include:

Table 2: Categorisation of theoretical work on land reform: Mechanisms and Direct Outcomes

	Increased (expected) security	**Individualization**	**Facilitation of transfer**
Investment	*Increases expected time horizon / reduces risk* Reduces security enhancing actions (de Meza and Gould, 1992; *Goldstein and Udry 2008;* Sjaastad and Bromley, 1997) Reduced effort of tenant farmers (Banerjee and Ghatak 2004, Besley and Ghatak 2009)	*Removes moral hazard (Alchian and Demsetz, 1972)* *Removes `Tragedy of commons' / free-riding - (Hardin, 1968) and under-utilisation of `anti-commons' (Buchanan and Yoon 2000)* Reduces economies of scale Use of low-risk low-return crops/technology	*Increases expected time horizon / reduces risk*
Credit	*Can be used as collateral (de Soto, 2000)* *Increase in demand for credit*		*Can be used as collateral (de Soto, 2000)*
Transfer of effective rights	*Increased leasing out (Conning & Robinson, 2007; Macours et al. 2010)*		*Consensual ownership changes to those with relative use advantage (i.e. better information, economies of scale, lower transaction cost) or* those looking to store value and those less risk-averse (Deininger and Feder, 2001) *Used as a liquid asset*
Time allocation and migration	*Reduction of security enhancing actions (Field 2007)*	Contracting problems may encourage self-use	*Can be sold / rented out by landowners*
Conflict	*Reduces potential returns to conflict*	*Reduces previous ambiguity of rights*	*Allows for transfer as a conflict resolution device* Increases possibility of contested transfer

Notes: Papers are classified according to the aspect of land right considered and the main direct outcomes the paper analyses. Impacts that are `positive' (i.e. roughly equivalent to welfare enhancing) are displayed in *italics*, and impacts that are negative are underlined. Impacts that could go either way are standard

(vi) Agricultural productivity: yields, technical and allocative efficiency

Productivity gains may result from the increased investment as well as through the transfer of land rights to other parties. For example, an increase in land ownership by those less risk-averse may result in more crops with `high-risk, high-return' profiles. At the same time, there is some risk that agricultural productivity may decline – for example, if land is transferred to owners who are using it mainly as a store of value, or if individualization increases the risk aversion of farmers.

(vii) Household consumption, income levels and stability

For land rights holders, the largest impact on household consumption levels will probably come about through changes in agricultural productivity. Note however, that non-farm income may also be affected, particularly if there are effects on labor use, resulting for instance from new investments on land and changes in used agricultural technologies. In the short-term, any observed increase in investment may come about through decreased consumption if households are credit constrained. In terms of consumption stability, greater access to credit and the ability to use land as a liquid asset may improve stability. Income fluctuations may, on the other hand, be greater if the intervention results in the adoption of riskier technologies and crops. Individualization of land may also reduce risk-sharing amongst the group. Finally, changes in land rights may alter the proportion of income received by various members within the household, such that it may be interesting to measure intrahousehold income allocation.

(viii) Land values and asset ownership

Increases in owners' rights towards land should increase its value, whilst increases in renters' rights may have the opposite effect. If the intervention means that land can now be used as a liquid asset, this may change households' overall asset portfolio. For example, ownership of land may increase while the ownership of other liquid assets decreases.

(ix) Food security

At a household level, the impact of land administration interventions on food security is likely to be closely linked to the impacts on income and consumption. At a more aggregate level, total production of food is likely to follow changes in agricultural productivity. An important exception however may be in the case where land is moved away from food crops (for example, due to lower risk aversion).

2.3.3 Municipal and higher-level final outcomes

Beyond the household-level impacts, there a number of important municipal and higher-level impacts that may result from land administrative programs. These include the following:

(x) Political support

The inclusion of this variable is mainly driven by the recent empirical literature suggesting that land reforms may have a significant impact on political preferences. Possible mechanisms may include lower dependence on local elites, support for the party that led the intervention or a greater participation in the market economy.

(xi) Increased tax base

A more accurate and detailed cadastral and registry system will increase the ability of a government to tax land. Moreover, citizens may be more supportive in paying such a tax if they believe that the government is supporting their land rights.

(xii) Land use planning

Particularly in urban settings, land planning provides a number of important benefits, and is likely to be facilitated by clearer land rights. One example is the provision of infrastructure, for which provision to insecure plots can be problematic. This may also apply to a certain extent in rural areas in cases such as the provision of irrigation schemes.

(xiii) Natural resource protection

Though not typically considered as investment, similar mechanisms to those outlined above will work for non-depletion of existing assets. However, there may be a concern that individualization of land will disrupt existing mechanisms to preserve natural resources that are common to the group. Moreover, increasing use of inputs such as fertilizers may have the side-product of increasing the pollution of water resources.

While the first important step is to define the range of outcomes that is expected to be affected, the related next step will be to translate these into specific measurable indicators, and to plan on the data collection needed to obtain information on all of them. Determining indicators that correspond to these expected outcomes is often not a trivial task, in part because they might differ depending on the specific context of the project. We will return to this issue when discussing data collection strategies and needs.

2.4 Heterogeneity of impacts and complementarities with other interventions

The theoretical models discussed above describe a range of possible impacts of land administration interventions. However, the empirical literature has shown that each of these impacts is not necessarily consistently found across interventions, even when carefully distinguishing how property rights are affected. This points to the fact that the standard models make a range of assumptions about the context from which the hypotheses on the link between land rights and outcomes are derived. Yet the assumptions are not necessarily relevant for all contexts. For instance, credit rationing may prevent both the increase in credit and the increase in investment predicted. Once such assumptions are relaxed in the theoretical models, they point to clear and rather intuitive predictions on the potential for heterogeneous impacts along a variety of dimensions. A well thought out theory of change should incorporate this, and doing so requires a good a priori understanding of the context, in order to evaluate the possible violations of the standard assumptions and consequent heterogeneity that could occur in any given project.

By exploring heterogeneity in impacts across beneficiaries of a project, impact evaluations can shed light on the relevance of context, as a given land administration intervention is likely to have impacts that vary across plots and households. For example, Ali et al. (2011) find that the impact of the intervention they study is greater for female-headed households, whose previous land rights were likely to be the most insecure. Exploiting such heterogeneity can give insights into the mechanism at work. And even if the change in rights is uniform across plots and households, there may be variation in impacts due to the necessary conditions for the mechanism to operate. For instance, Carter and Olinto (2003) show that the total investment impact will be greater for wealthier individuals in the presence of credit constraints.

While all too often heterogeneity is considered as an afterthought, as a way to further understand initial results, they should in fact be carefully considered ex-ante as part of the causal chain. Indeed, in order to credibly and rigorously evaluate impact heterogeneity it should be planned for ex-ante, so that sampling design and data collection can be organized in order to shed light on the related hypotheses. The dimensions along which one expects heterogeneity to occur need to be defined, and variables that capture those dimensions (gender, wealth, credit access, etc.) need to be incorporated in survey design. Moreover, sampling procedures will need to be adjusted in order to assure enough empirical variation among those dimensions, and in

order to avoid ex-post cherry picking of interesting heterogeneity results. We will return to this point when discussing sampling challenges below (see 3.6).

In some cases, the presence of other binding constraints may even suggest the potential need to complement land administration interventions with other complementary interventions. For instance, in certain contexts, the potential for land administration projects to affect productivity might be limited by lack of access to new technologies or credit. When such complementary interventions can become incorporated in the overall project, or even when they are envisioned to occur in the same region and target population, the possibility exists for evaluations to shed light on the possible complementarity of these interventions. While this can be very promising, further methodological challenges arise from the need to establish causal inferences regarding both the land administration and these complementary interventions, which we address in section 3.5.

2.5 Spillover effects

The discussion so far has focused on the impact of land administration interventions on households and areas targeted by the intervention. In addition to these direct impacts, however, there are also likely to be spillover effects in areas not subject to the intervention. These spillover effects are important to consider when designing an evaluation for two major reasons. First and most directly, the spillover effects may in themselves be important, and therefore count as an impact of the intervention that requires evaluation. Second, if the spillover effects impact upon a group that is considered as a `control'- and hence implicitly assumed not to have been impacted by the program – this may complicate attempts to measure the impact of the intervention on those targeted.

In the case of land administration interventions, there are several possibilities for spillover effects. One group of such effects is likely to be the result of anticipation amongst households not targeted. Since land administration interventions are generally sanctioned by the national government, it is very reasonable for non-targeted households to believe that they will be targeted in the near future. This belief is of particular concern for evaluating land administration interventions due to the importance of *expected* security in the mechanisms outlined above. An intervention such as land-titling may, for example, increase expected security even amongst non-titled households if they believe that in the near-future they will receive such a title. On the other

hand, for certain households, expected security may decrease if they believe that there is a high probability someone else may receive the title. The two other channels discussed above - individualization and facilitation of transfers - may also potentially suffer from anticipation effects. If it is believed that an individual rather than a group will soon hold rights over a plot, group members behavior is likely to change in a variety of ways. Meanwhile, if it is believed that transfers will soon be facilitated, this may temporarily reduce land transfers, or households may change the way they use the land if they anticipate a future sale.

Another set of spillovers that may occur are those that result from the relationship between land markets across both targeted and non-targeted areas. If the intervention increases the probability of land being leased, those that lease the land may come from outside the targeted area. An increase in the value of land that benefited from the intervention may result in a decrease in the value of non-targeted land. Migration and labor supply decisions may also lead to spillover effects if they are large enough. This set of spillovers should be noted in particular when measuring the impact of interventions on land transactions, since such transactions can easily involve households from outside of the treatment group.

3. Evaluation Design and Methods

In spite of strong theoretical predictions on the channels and directions through which land rights could affect agricultural and welfare outcomes, the empirical evidence on the potential benefits from titling and other land administration programs remains inconclusive. There is mixed evidence of effects on investments and land allocation and, if anything, zero impacts found on credit. We now return to the second possible reason for this mixed evidence on the impacts of property rights insecurity and land rights interventions: it could very well be related to methodological weaknesses. Indeed, a lot of the existing "evidence" comes from evaluations or empirical studies that are likely to suffer from severe endogeneity bias.

A first important explanation for this record is that much of the empirical evidence is based on associational-based evidence from observational studies. These tend to investigate the correlations at a given point in time between the distribution of land rights – the result of past transactions, institutional changes and policy interventions - and individual outcomes. In such observational studies, interpreting the relationship between land rights and outcomes as the causal effect of a specific policy change relies on strong assumptions, as many unobserved

confounding factors could drive the observed correlations. For instance, landowners with formal titles usually differ in many ways from those without such documents, so that attributing their different behaviors and outcomes to their land ownership status is simply not credible. This makes it crucial to evaluate the impacts of specific policy interventions.

When designing an impact evaluation for land tenure interventions, selectivity into treatment is the main challenge to be addressed. Because land tenure interventions tend to benefit and affect the rights of specific sub-groups of individuals, e.g. those with initially more insecure tenure or farmers in regions with a higher agricultural potential, the potential outcomes of beneficiaries with or without the intervention are likely to differ. Moreover some confounding factors correlated with the participation to the intervention (e.g. different weather conditions if the two groups are not in the same area) could drive the differences in the outcomes of the treated and non-treated groups. Simple comparisons of the outcomes of beneficiaries with those of non-beneficiaries are thus unlikely to produce unbiased estimates of the intervention's impacts.

Rigorous evidence can only be produced if the impact evaluations are designed in ways that carefully account for the allocation of treatment and allow identifying a comparison group that provides a valid counterfactual for the outcome of program participants in the absence of the intervention. This can be done by using experimental (i.e. randomized assignments) or non-experimental methods. Experimental methods require fewer assumptions, often provide a clean causal interpretation, but can be challenging to implement. For non-experimental methods, more assumptions will need to be made, and hence evaluations will require careful checking and the provision of evidence in support of the validity of these assumptions.

In either case, the data collection strategy is at the core of an impact evaluation. The data that is collected and the data analysis come together, but data collection is the first and the most demanding stage, and will determine the subsequent data analysis. Moreover, data collection needs to be built into the early program operations, as, for instance, an experimental evaluation will require incorporating some beneficiaries (at an individual or a more aggregated level) at different dates based on lotteries, or a non-experimental matching-based evaluation will require identifying potential beneficiaries in areas which are not included (at least in the early years) in the program.

Prospective Impact Evaluation (IE) plans are hence strongly recommended for obtaining a valid comparison group, and thus credible estimates of the program impacts, and using efficient data collection processes. Those plans should include the definition of a comparison group, the data to be collected at different stages of the intervention, and the way the data will be analyzed.

We now discuss the different methodological options; for each alternative, after presenting the strategy for identifying the impacts of an intervention, we discuss the implications for data collection.

3.1 Identification based on randomized assignment

The most rigorous, and in some senses the most straightforward, way to assure that one can identify the causal impact of a land administration intervention is to assign the intervention randomly among a large group of villages or individuals that is eligible for the intervention. By comparing the mean outcomes among those that randomly were assigned treatment, with the mean outcomes of those that randomly were not allocated the intervention, the causal impact can often be measured without having to worry about any biases or confounding factors. By making sure that assignment to the treatment group is independent from potential outcomes, randomized controlled trials (RCT) provide an adequate comparison group to the group of beneficiaries, i.e. a group from which one can learn about the potential outcomes of beneficiaries had they not been treated. In general, RCTs therefore provide the most reliable evidence on the causal effects of interventions. The general advantages of RCTs have been discussed in much detail elsewhere (e.g. Duflo, Glennerster and Kremer, 2008) and certainly hold for the case of land administration interventions.

However, given the sensitivity of land rights in many contexts, political will is key to implementing a rigorous RCT of a land rights intervention. Moreover, the use of RCTs for such evaluations raises several issues regarding implementation modalities and statistical power, internal validity issues, and external validity considerations. Evaluation designs therefore ought to carefully consider these political, logistical, practical and statistical factors from very early on in the project cycle.

First consider implementation aspects. There are several ways to implement RCTs, and notably different possible units of randomization, and those have bearings for both the measurement of impacts and the operations. On the one hand, because land interventions, in

particular those that seek to clarify rights, involve activities that are performed at the level of communities (such as information campaigns, surveying, conflict resolution), and because the externalities they generate between neighbors are likely to be strong (due for instance to changes in local land conflicts, anticipation of future eligibility, or land markets equilibrium), randomization generally must be conducted at the level of some sort of geographic cluster (which could be localities or communes).[5] On the other hand, to insure the balance of characteristics between the two treatment groups and achieve a given statistical power, randomization needs to be performed at the level of sufficiently numerous (and thus small) areas. Typically, as soon as outcomes are somewhat correlated within clusters, RCTs with less than one or two hundred clusters are likely to be underpowered. Thus, for land interventions with clarification activities, a design that accounts for both local externalities and statistical power could consist of several hundreds of small geographic clusters.

However, such RCTs put some constraints on operations. In most cases, it seems difficult to implement a land intervention in dispersed small geographical clusters and not in other neighboring clusters. The typical solution is therefore to implement the RCT through in the context of a staggered phase-in where control clusters would be incorporated in a second phase. Operations would then have to be adjusted to accommodate the RCT. In particular, the surveying and clarification operations that involve a pre-cadastral sweep of covered clusters have to be adapted to involve at least two sweeps. A strictly enforced timing of the intervention is thus critical. That will be easier to do when the program administration is centralized so that the schedule of phase-in is controlled, but probably harder to do if more (local) actors involved. Also the constraints on the phase-in schedule will incur some operational costs (such as more transportation) that have to be taken into account in the budget proposal laid out in the impact evaluation plan.

We are aware of only two RCTs having been implemented for evaluating the impacts of a land tenure intervention.[6] The first is an ongoing evaluation of a pilot land surveying and certification program ("Plans Fonciers Ruraux") implemented by the government of Benin with support from the MCC. A preliminary impact evaluation was conducted by Selod et al. (2012); it

[5] This might be less necessary for interventions that only imply plot-level formalization/registration activities, in which case the treatment of individuals, with neighbours having their rights formalized later on, would induce less externalities and thus be feasible.

[6] Other related RCTs have been planned but failed to be implemented for a variety of reasons, while some other ones are ongoing but, to our knowledge, do not yet have results (Rwanda, Mongolia).

rests on the random selection of about 300 villages that benefited from the program. The randomization was conducted at the village level within each commune (the control group should benefit from the program when it will be scaled-up nationally). The possible occurrence and effects of some spillovers of the intervention in control villages need to be documented (see below), but this set-up should otherwise warrant sufficient statistical power.

Also in an African country, but now in a urban setting, Ayalew, Collin, Deininger, Dercon, Sandefur and Zeitlin (2011) have implemented an RCT for evaluating the variation of a titling component of a tenure securization program. The RCT was run in two urban slums in Dar es Salaam, Tanzania, and consisted first of providing access to formal land titles to informal settlers at randomized prices and second of offering additional price discounts conditional on designating a woman as owner or co-owner of the household's land (the overall program also had activities for clarifying rights). The randomization into the treatment group was conducted at the level of 'blocks' (contiguous groups of approximately 40 parcels), while the randomizations of the offered prices and discounts were at the individual level. There were only about 20 blocks but about 1500 individuals.

To our knowledge, no RCT has yet been conducted for evaluating land tenure interventions in LAC countries. In section 4, we return to a planned randomized evaluation in Haiti. Additionally, Gonzalez, Corral, Simning and Winters (2011) propose a plan for evaluating the impacts of a land surveying and titling program to be implemented in Paraguay. It is based on a randomized phase-in of eight districts, four of which will be incorporated initially and four two years later. Given that the number of randomized units is small, the balancing of characteristics of the treatment and control groups might not be achieved and statistical power is likely to be limited. The proposed strategy thus relies also on complementary unconfoundedness and difference-in-difference identification approaches, using the baseline data to control for observables and pre-program outcomes.

The second set of issues to take into account relate to internal validity. In spite of the strong comparability of the treatment and control groups, RCTs are not always free of biases. Because some of the program components and effects are likely to affect behaviors of households and farmers whatever the cluster they live in, it is important to be able to distinguish the direct effects of the surveying and/or formalization of the plots owned by individuals in the treatment group from broader program and spillovers effects. Broader program effects are likely to occur in

particular if public awareness and information campaigns on the importance of secure lands rights and responsibilities of land owners and occupants that precede the pre-cadastral sweep will cover the entire pilot communes. Furthermore, the clarification or registration of rights process in some treatment localities might affect overall perceptions and anticipation of future tenure security in yet uncovered clusters. Investments incentives anticipating the intervention, local conflictuality and the prices of land market are different channels of such spatial spillovers. As little evidence is available on those, disentangling direct effects from indirect and/or spillover effects would be of interest. A planned RCT evaluation in Haiti (see section 4) addresses this for instance by collecting data among a non-experimental comparison group that would serve for estimating the extent of spillovers on the control clusters. More general, from an implementation perspective it is useful to plan ex-ante on how to avoid, to the extent possible, that those that will benefit only after the phase-in of the control group do not modify their behavior in anticipation of future benefits. Managing expectations and clear communication by the program team (on timing and likelihood of the phase-in) can help homogenize expectations in the control group, making interpretation of differentials with the treatment group easier to interpret.

The third set of considerations relate to the external validity of RCTs. Some concerns are shared with other evaluation approaches. The specificities of the areas selected for the evaluation might limit the external validity of the results, as applies to any evaluation of a small-scale program. When other interventions are conducted in the same areas (for instance for agricultural development), it is important to seek to isolate the impacts of each intervention, e.g. using orthogonal randomizations of each, to understand their complementarity and the effects that could be expected from scaling-up the interventions together or separately (more on this below). An external validity concern more specific to RCTs is the length of the experiment: the effects of land interventions can take time to appear, so that it is important to observe the outcomes of the treatment and control groups after a sufficiently long period of time. A staggered phase-in might put limits on the time before the control clusters are incorporated, but two to three years, depending on the context, seems a minimum to observe impacts on some investments (e.g. land improvements or tree planting) or income (e.g. perennial cultivations).

Fourth, unexpected political developments might complicate and possible endanger compliance with the experimental design. Given the political sensitive nature of land rights, it might be hard in the first place to convince government counterparts to agree on a randomized

allocation. Moreover, once the intervention has started, local pressures might increase and demands by households in the control areas possibly could mean that the experimental control group also receives land rights, leading to contamination of the experimental design. On the other hand delays, logistical, administrative or political problems could imply that part of the treatment group does not receive the land rights in time. To avoid such complications, it is key to have strong buy-in from the different partners, and to build in the randomized roll-out into the operational procedures. That said, to the extent that non-compliance is limited, the randomized assignment can still allow the identification of causal impacts by using the randomized assignment as an instrument. This will generally have less power, but provides a second-best outcome and some insurance.

In a randomized experiment, the sample to be surveyed is to be determined concomitantly to the experimental design. Several points are worth noting here. First, statistical power will depend mostly on the number of randomization units as discussed above. Once there are a sufficiently large number of those, as discussed above, cost-efficient sampling procedures will consist of drawing small numbers of households or farmers in each of the randomized units. However, attrition and incomplete compliance with eligibility have to be accounted for in power calculations. Second, the baseline survey can be useful for checking that the randomization was conducted properly, for improving the precision of the statistical estimates of the impacts (and reduce the size of the samples to be surveyed at follow-up), for investigating the heterogeneities in impacts and assessing the external validity of the impact evaluation, and finally for validating data collection procedures. Third, when some program components can have spillover effects and contaminate the control group, a complementary non-experimental comparison group drawn in neighboring areas out of the program can allow quantifying those effects; sample size has to be increased accordingly.

3.1.1 Encouragement designs

In case there are program components that in theory could affect the whole population, but in practice might have limited impact without additional complementary interventions, the randomized addition of such complementary interventions can help to evaluate their impact. For example, it would be possible to randomize information about a certain legislative change that increases tenure security to analyze the impact of increased security. Similarly, it would be

possible to randomize subsidies for a titling program that implies cost for individual households in such a way that increases (randomly) demand by households for such titles, thereby allowing an evaluation of the impact of titles. The evaluation in Tanzania referred to above uses such a design. Such evaluation designs are referred to as encouragement designs. While they allow establishing causality, they only show impacts for the population of people that change behavior because of the encouragement, and hence only allow estimating a local average treatment effect. In addition, they may have low statistical power, as take-up among the "encouraged" population might be low. As such, they are not a first best strategy, but they should be considered as a possible option for components or interventions that because of their large scale or reach do not allow identifying another plausible counterfactual.

3.2.Non-experimental approaches

Using experimental approaches to evaluate impact may not be possible due to a number of practical or political considerations. For example, certain infrastructure projects are most cost effective if implemented systematically by location instead of randomly. Some projects may also be universal and cover an entire population. If experimental approaches are not possible, there are a number of alternative non-experimental approaches that can be used. A few of the key possibilities are discussed in this section.

3.2.1 Regression Discontinuity Design

In many cases, a second-best approach for obtaining credible estimates of impacts of titling interventions would be based on a regression discontinuity (RD) design. Often, when studying the effects of a policy intervention, eligibility to the program is assigned based on transparent rules, with explicit criteria and clear cut-off values, rather than on discretion of program administration. In such cases, comparisons of individuals that have very similar characteristics but lie on different sides of the eligibility cut-off can provide reliable estimates of the program impacts for the specific population of individuals "close to" the cut-off.

The RD design, now widely used in applied economics, can be applied for obtaining estimates of the impacts of land administration programs in several ways, depending on program implementation. For example, the targeting of land interventions might rest on some explicit criteria that are effectively enforced and generate a discontinuity in treatment assignment that

does not correspond to any substantial differences between the two groups. For instance, some titling program might target smallholders cultivating parcels of a size below a given threshold, so that farmers with slightly larger parcels are excluded. In designing prospective impact evaluations, it is hence important to understand how the program will be targeted, based on which data eligibility will be determined, and whether there exist an eligibility rule for which such a threshold can be found. Indeed, when working with the operational teams prior to finalizing the program design, it can be helpful to introduce such thresholds specifically for the evaluation, for instance in making intended targeting rules more precise and in assuring that data will be systematically collected to apply those targeting rules. This then can allow identifying those above and below the threshold.[7] In such cases, equally important would be to guarantee that potential beneficiaries might not be able to manipulate the information that determines whether they are below or above the threshold (see below).

Existing examples of studies that use RD design to study land administration projects are not necessarily based on such a prospective design, but are nevertheless interesting in providing other ideas regarding this approach. One example includes land surveying and titling interventions that are implemented at a small scale, often for their piloting, usually cover areas delimited by precise borders, such as a few communes or municipalities. One can then compare the outcomes of individuals owning or exploiting parcels lying on the two sides of the borders; this approach was followed by Ali, Deininger and Goldstein (2011) to estimate the impacts of a pilot land titling program in Rwanda.[8] One concern with this type of discontinuity is that the borders of the selected pilot areas could correspond to specific geographical barriers (river or mountain range) that could be associated with changes in some determinants of agricultural production (such as climatic or soil conditions). Externalities on the control group could also be a concern if households beyond the border learn about the intervention and anticipate it in some way (more on externalities below).

[7] For example, if the objective of a program is to target "poor smallholders", it would help to come up with specific observable threshold that help determine who is poor (e.g. based on an asset index or a proxy means test on census type data) and a threshold farm size for smallholders.

[8] In Rwanda, a large-scale land tenure regularization program was introduced to map, clarify and register rights (a formal certificate is then obtained upon payment of a fee). The study examines the impact of its pilot, implemented in four rural areas (with about 15,000 parcels). The data is from a survey conducted in April-May 2010 among about 3,500 households within 125 to 1250m on both sides of the boundaries of the four pilot areas. Tenure regularization was compulsory so that all private parcels in the pilot areas were treated. The impacts of the pilot program are estimated in a regression framework by the difference in average outcomes of treated and control households (without an extrapolation) but also introduce spatial fixed effects (for specific small radiuses).

Another potential application of RD can be when legal rules might generate thresholds determining which, and in what ways, different parcels or individuals are affected by the intervention. Vranken et al (2011) thus consider the effects of the restitution of land to former owners and their heirs at liquidation of former communist cooperatives and state farms in the early 1990s in Bulgaria. They exploit the discontinuity generated by a law preventing excessive land fragmentation, which had plots below a given size (0.3 acres) remain undivided in co-ownership among the different heirs.

Let us consider the case where treatment is decided based on the value of a given assignment variable (such as location with respect to the border of treated geographical areas or size of owned parcels). The key identifying assumptions in the RD design approach is that the potential outcomes of individuals under both treatment and non-treatment conditional on the assignment variable are continuous in the value of this variable. In other words, there should be no jumps in the potential benefits or costs associated with treatment at the treatment threshold of the assignment criteria. This guarantees the comparability of individuals lying just above and just below the cutoff, and allows estimating the effects of assignment to treatment by the difference in the outcome. This requires in particular that individuals cannot manipulate the variable determining assignment, as individuals who would benefit most from the intervention would otherwise tend to be more often above the threshold.

It is therefore important to analyze, once data is collected on the eligibility criteria, that the distribution of the key variables around the threshold is smooth, and does not show any signs of manipulation. In particular, if the distribution has much higher density just below than just above the threshold, this indicates there might have been manipulation, in which case one should not rely on the RD for identification. Hence this needs to be analyzed as soon as possible, if RD is being considered as an approach. Further tests can also be performed (both with data used for eligibility and with data collected for the impact evaluation) to establish empirical support for the identifying assumption. In particular, one would want to show that there are no discontinuities in average covariates at cutoff, as that would indicate that individuals on the two sides of the thresholds do not compare well and that the continuity of potential payoffs does not hold. On the other hand, it also helps to analyze possible discontinuities in the average outcomes at other values of the assignment variable, which would suggest the presence of jumps that are unrelated to the intervention.

If the identifying assumptions seem reasonable, to obtain an estimate of the impact, one estimates the difference in the values taken by two regression functions estimated on each sides of the assignment cutoff. It requires a local extrapolation, which is obtained with local linear regressions estimated on some bandwidth (see Imbens and Lemieux 2008 or Lee and Lemieux 2010 for details). Sensitivity to the specification choices, such as the used bandwidth, need to be performed.

An important limitation of RD design estimates is their limited external validity. Indeed the effects of assignment to the intervention are estimated only locally around the threshold, i.e. for individuals that may have specific characteristics and do not compare well to other potential beneficiaries. For instance, estimates exploiting a geographical threshold would inform on the impact of an intervention for individuals with parcels of land near the border. Similarly estimates exploiting some targeting rules inform on the impacts for groups that have characteristics close to the assignment cutoff. One can then document the extent to which those groups look similar or differ from other potential beneficiaries, but the RD design cannot provide causal estimates of the impacts for the later group. However, for individuals around the admission cut-off (e.g. near the border in a RD design based on geography) allocation into treatment can be considered as random and a RD design provides credible estimates of the impacts of a land rights intervention.

The sampling of data collection needs to be adapted in an evaluation based on a RD design. First, in this case the sample is to be constituted mostly of individuals with characteristics such that they lie near the treatment cut-off. For instance, in a RD design evaluation based on geography and distance to a given treatment border, the households in the main sample would be drawn from the population of households living (or exploiting parcels) within a given narrow distance on the two sides of the border. If the RD design evaluation was based on a means test, or another eligibility score, the sample would be constituted of households with scores close just below and above the cut-off. Second, in order to assess the external validity of the estimates (see below), the sample would have to be complemented by a set of other beneficiaries. However this later sample will not serve for identifying the impacts of the intervention, and would just be of limited size.

3.2.2 Evaluations without baseline data: cross-sectional regression and matching studies

The main alternative non-experimental method consists in using comparison groups of non-beneficiaries who have similar (or sufficiently close) observable characteristics to the ones of beneficiaries of the land rights intervention. The program impacts are then obtained using regression or matching estimation econometric techniques. [9] As these evaluations rest on more assumptions, which by definition can not be tested, these provide much less credible estimates of the intervention impacts. For assessing the validity of the comparison group and improving the quality of the estimates, it is critical to use both pre- and post-intervention data. Hence this type of evaluation, and the associated data collection, also has to be planned before the intervention.

Before turning to this more favorable setting with baseline data (in the next subsection), we discuss the case in which only post-intervention data is available, typically when the impact evaluation is decided and conducted after the program has begun to be implemented. We discuss this approach for completeness and expositional purpose, but strongly recommend against it given the strong assumptions needed to interpret results as causal. But if anything, such an approach may be helpful for the ex-ante evaluation based on observational data. For ex-post evaluation of program impacts, this approach should only be used as a last resort when the evaluation was decided after the program was implemented.

In this case, the program impacts can be identified only by comparing beneficiaries to some non-beneficiaries with similar characteristics. The key identifying assumption here is unconfoundedness (or selection on observables), i.e. that conditional on a set of observables, an individual's potential outcomes with higher or lower degrees of security are similar to those who possess such higher security and those who don't. In the case of a titling intervention, unconfoundedness would occur if the likelihood of receiving a title is independent of individuals' potential outcomes. This would require that the allocation of titles depends on some cost factor that is independent of beneficiaries' characteristics and potential outcomes. Similarly, in studies of the relationship between owned rights and investments, the assumption requires that

[9] Many studies have examined the link between tenure security or the benefit from a land property rights reform and farmers' behaviors and outcomes, such as investments or access to credit, using regression analysis with cross-sectional data. For instance, Fenske (2011) reviews 54 empirical studies that have tested the link between land rights and investments using African data at the household (or plot) level in a regression framework. There are also a number of similar studies in Latin American settings. For instance, Bandiera (2007) analyses the effects of land tenure, *i.e.* whether the farmer is the owner of the land or a tenant, on crop choice. Deininger and Chamorro (2004) use regressions with household fixed effects to examine the effects of a titling program implemented in Nicaragua.

individuals who do not enjoy more secure rights would benefit as much from those rights as individuals who do possess them. Unconfoundedness is not testable and its violation can bias the parameter estimates in an unforeseen direction.

There are several reasons why unconfoundedness may not hold. First, whether people request titles/get titles, often depend on individuals' potential benefits from more security or from perceptions on insecurity, which are unobservable. Second, some characteristics of plots and cultivators, that are correlated with tenure security and may affect investments, are particularly likely to remain unobservable when one compares differences across geographical areas with different farm (e.g. soil fertility) or household (e.g. farmer's human capital and ability, family composition) resources.

For reducing the concerns above, studies have relied on comparisons between farmers in small geographic areas (using spatial fixed effects), plots within the same households (using household fixed effects, as in Goldstein and Udry, 2008) or cultivated by the same farmer (using farmer fixed effects). However nearby farmers or plots might still differ. For instance, while spatial fixed effects could control for some unobserved heterogeneity such as ecological conditions, others such as the history of market transactions would probably not be removed. Similarly, even with household fixed effects, the process of allocating plots across members of a single household could drive some empirical patterns. And, at the farmer level, the selection of which plots to title is likely associated with differences across plots. Thus, in most settings, it is unlikely that unconfoundedness holds. In addition, reverse causality with some investments performed to increase tenure security would generate a similar empirical relationship, so that the direction of causality could be uncertain.

Although unconfoundedness is ultimately not testable, its plausibility can be assessed, and this should certainly be done whenever possible. A first approach for this consists of estimating the effect of the treatment on a pseudo outcome. For instance, one could use information on lagged measures of the outcome (e.g. investments or productivity) before receipt of a title, and test for the presence of pre-treatment differences conditional on the other covariates, but this would require some type of baseline data.[10] A second approach consists of estimating the effect of a pseudo-treatment on outcome, for instance when subsequent groupings of geographic areas

[10] Note that one would need to test for pre-existing differences in the distribution of the pseudo-outcome, not only its mean.

are incorporated each year in a program, one could compare the outcomes before treatment of individuals in two groupings of areas that have not yet been incorporated.

Regression studies also rest on an overlap assumption, requiring that for all values of observables, the probability of treatment – receiving a title or enjoying more secure rights – is more than zero and less than one. This assumption is in principle testable, as one can estimate the propensity score for treatment conditional on observables. Lack of overlap in covariate distributions makes conventional estimators sensitive to minor changes in specification and leads to imprecise estimates. Indeed, the estimates will then rely a lot on extrapolations based on some specification assumptions, and slight changes in those assumptions could make the estimates vary considerably. In this case, the average effect of treatment can be estimated more robustly by improving the balance of the sample using matching techniques. However, finding good matches for a sufficiently large share of beneficiaries may require a lot of data.

Matching is usually performed on a propensity score (the predicted probability of treatment given the value taken by observables), so that each treated individual is matched with the closest control (without replacement).[11] Some tests confirm that matching estimates can perform much better than regression estimates on non-experimental data (Imbens and Wooldridge 2009). That said, while matching works well when program participation is based on selection on observables, this is arguably unlikely to be the case for titling, for which unobserved demand might play a key role. That is possibly why matching has not been used very often for studying the effects of tenure rights though, though we found a few applications (Bandiera, 2007; De Laiglesia, 2004, Zegarra et al. 2007).

3.2.3 Difference in difference evaluations

We now return to the recommended practice of preparing the IE before program implementation, as should be the rule. When a randomized experiment is not feasible and a regression discontinuity evaluation is not feasible either and/or likely to have a too limited external validity, there remains some scope for evaluating the impacts of a program, in a difference-in-difference (DiD) scheme, possibly combined with matching, when an intervention is phased-in sequentially and the impact evaluation is planned sufficiently in advance, so that baseline data can be

[11] A related method consists of trimming the sample by removing observations with values of propensity scores that lie to close to zero or one.

collected among some program participants before they benefit from the intervention. However, these evaluations require surveying large samples, and ideally include data from multiple survey rounds, and therefore tend to be costly.

When the intervention has been phased-in in a sequential way and introduced in different areas in different years, which is usually the case of large-scale interventions, the groups of beneficiaries treated later can serve as a comparison group for those treated earlier. The concern is that the comparison of different groups at a given date is potentially biased because, as the intervention was not incorporated randomly, the groups may differ in unobserved ways (e.g. areas with a higher agricultural potential may have been covered first). In the same way, the comparison of the same groups at different dates is potentially biased because of other changes that have occurred over time.

In this case, DiD methods that compare the change in the group that was treated in the second period to the change in the group that was not (or not yet) treated in the second period can under some assumptions eliminate both biases.

In a DiD setting, the effects of the intervention can be estimated in a standard regression model with an interaction term between the observation period and assignment to the treatment and control (or late treatment) group, or in a panel model with fixed effects when the data follow observational units (plot, individual, household, etc.) across time. One can incorporate controls for a set of observables. The key assumption here is one of common trends - econometrically this requires that residual variations in the outcomes be uncorrelated with treatment (and observables) -, so that any unobserved factors that affect changes in outcomes would affect in the same ways the changes in outcomes of treatment and control individuals. This assumption is related to the one in models assuming unconfoundedness, except that it now controls for time invariant unobserved heterogeneity between the two groups. It remains strong however, and any unobserved heterogeneity that is not time invariant would lead to different trends and hence bias the results. For instance, suppose that land titles have been delivered first in poorer areas and that some catching-up would be taking place independently of the intervention. In this case, farmers' productivity and income would increase faster in those areas, and this would bias the estimates of the effects of titling. Variations in climatic conditions could similarly drive differences in trends. The DiD approach can also be combined with matching estimates, which can in some cases increase the plausibility of the common trend assumption.

Several non-experimental studies of the effects of titling interventions have relied on DiD estimates. Field, Field and Torrero (2006) for instance investigate the effects of the Special Rural Cadastre and Land Titling (PETT) program, which was implemented in Peru starting in 1993 and consisted in a complete securization process with surveying and titling of parcels and establishment of a cadastre in rural areas. Zegarra et al (2007) combine matching with DiD estimates for the same program. And de Janvry, Emerick, Gonzalez-Navarro, Sadoulet (2012) examine the effects on migration of Procede, a large-scale land certification program implemented in Mexico from 1993 to 2006.

As for studies based on cross-sectional regressions, the identification assumption is not testable, but its plausibility can and should be assessed. One approach for this consists of testing for the presence of pre-intervention trends (similar to a pseudo-outcome) when data is available for several points in time prior to treatment (e.g. de Janvry, Emerick, Gonzalez-Navarro, Sadoulet, 2012; de Janvry, Gonzalez-Navarro, Sadoulet, 2011). Another approach (similar to a pseudo-treatment) consists of testing for differences in outcome changes across different sub-groups of control individuals, e.g. two groups of areas that will be covered by a land titling intervention at two different future dates.

For implementing DiD evaluations, data need to be collected among beneficiaries and a comparison group, preferably at several dates, but at minimum before and after they benefit from the new more secure land rights. The sampling of the evaluation surveys will have to be particularly careful in this case. First, the comparison group can be selected among potential beneficiaries to be included in a later phase of the program, but in this case a sufficiently long lapse of time is needed before the later group is incorporated in order to estimate the medium or long-run impacts of the intervention. Second, when using DiD with matching, a very large sample of the comparison group needs to be collected, in order to obtain within that sample enough observations that can be identified as close matches to the treatment. Otherwise, as discussed in the preceding subsection, the estimates will rely on specification assumptions (for obtaining the counterfactual outcomes through extrapolations) and can be very imprecise. As the costs of collecting data among large samples might be prohibitively expensive, this is often an important consideration when deciding whether to opt for matching. One can also try to stratify the data collection (from the baseline onwards) on a number of key characteristics that allow matching households in the treatment and comparison areas. Third, the comparison group will be

likely poorer when it is drawn in geographical areas that are more remote from the treatment areas, or in different administrative divisions, as in this case it will be more likely that confounding factors drive some different trends in the two groups. These concerns will have to be balanced with those related to spillovers and contamination of the control group.

3.2.5 Natural experiments or non-experimental Instrumental Variables

A last type of evaluations, which consist of exploiting some sources of arbitrariness in the allocation of treatment, can be performed in some settings. However, those evaluations, mainly used ex-post by researchers, can hardly be planned in advance and thus do not fit well the needs of program managers.

Sometimes, whether some individuals are treated by a policy intervention depends on some exogenous factors that are independent from the potential benefits and costs treatment would incur for them. For instance, in the case of land titling programs, some arbitrary rules in program administration might determine which areas are treated first and which ones only later without an explicit targeting based on observable characteristics and related to potential outcomes. Castaneda Dower and Pfutze (2011) thus argue that, for the Mexican certification program Procede, the timing of the first contact made by program staff with Ejidos was determined without an explicit targeting strategy, but mainly depending on distance to the state capital where the staff were based (which they assume is not associated with potential outcomes of program beneficiaries). In other instances, whether the program is actually implemented or not depends on factors that are independent from individuals' choices and outcomes. Galiani and Schargrodsky (2010) study the case of squatters who occupied from 1981 urban land that they partitioned into small parcels; while the squatters believed the land belonged to the state, it was actually private property of 13 landowners. In 1984-1986, the Congress of the Province of Buenos Aires passed a law to expropriate these parcels and allocate them to the squatters with a monetary compensation to the former owners, but, while 8 owners immediately accepted the expropriation and associated compensation, the other 5 contested the decision in courts. As a result, some squatters obtained formal land titles in 1991 while others had to wait until the dispute was settled in 1998. The authors argue that the two groups of squatters were not different, e.g. all came from the same emigration region, and there was "simply no way for the occupants to know ex-ante, at the time of the occupation, which parcels of land had owners who would

accept the compensation and which parcels had owners who would dispute it". Other historical accidents, such as political changes affecting the content and implementation of interventions, might provide similar conditions for "natural experiments" whereby some exogenous factors influence the allocation into treatment.

Natural experiments allow the identification of the causal effects of interventions for subpopulations that have their treatment status modified by the exogenous factor. In the econometric analysis, these are the subgroups that are induced by the instrument to change the value of their endogenous treatment, and one estimates the local average treatment effects (LATE) for this subgroup. Two key identifying assumptions need to be satisfied. First, exposure to the exogenous factor determining treatment needs to be random (random assignment). Second, this exogenous factor (the instrument) must not directly affect the outcome (the exclusion restriction). For instance, in Galiani and Schargrodsky's setting, whether landowners contest the expropriation ruling (and the allocation of squatters across initial landowners) is arguably random, and in addition the appeal decision and legal dispute is assumed to affect outcomes only through the land rights that squatters own during the subsequent years (but not in other direct ways). These assumptions are not testable, but can be sometimes assessed using some of the principles we discussed, notably of pseudo-treatments. For instance, when the instrument affects a well-identified sub-population, one can check whether it is also correlated with changes in the outcomes among other groups. A third assumption is also necessary for interpreting the estimates, namely that the effects of the exogenous factor on program participation play in the same direction for all individuals.

Under these assumptions, the causal effects of the intervention can be obtained for the subgroup identified by the instrument. One limitation of natural experiments, which is similar to the one met in RD design analyses, is that these subgroups are not always the most interesting, but the data cannot inform about the effects of the intervention for other subpopulations without performing extrapolations that rely on strong assumptions. Little can be done to extend the limited external validity.

3.2.6 Synthetic control methods

In case an intervention (or a component) takes place at an aggregate level, such as a strengthening of local institutions or a change in local land legislation, and affects an aggregate

subgroup, such as a region (or set of regions), and aggregate/macro data (e.g. at the regional level) is available, some comparative case studies methods can be used.

These will, more specifically, require access to aggregate data on the outcomes expected to be impacted and characteristics of the treated aggregate units. For instance, for investigating the impact of a strengthening of land rights management institutions in one or a few regions, one could consider, when those are available, using aggregate measures of investments and agricultural productivity (or other intermediate or final outcomes such as conflicts) at the same regional level for different periods of time before and after the intervention. A comparative case study approach would then obtain a counterfactual of the aggregate change in outcomes using data from untreated aggregate units, usually by comparing the changes in the aggregate outcomes in the treated regions with those in other untreated regions. In general comparing outcomes across regions (or other aggregate units) is not very reliable as some confounding factors could drive the trajectories observed in the different areas.

For reducing this concern, the synthetic control method (Abadie, Diamond, Hainmueller 2010) proposes to construct a more adequate counterfactual as a weighted sum (the synthetic cohort) of the outcomes of the different untreated regions (or other aggregate units), the weights being obtained by making the synthetic cohort as comparable as possible, in terms of both outcomes before treatment and observable characteristics, to the treated aggregate unit(s). The impacts are then measured as the difference after treatment between the outcomes measured for the treated unit(s) and synthetic cohort (the counterfactual), and statistical inference can be obtained by investigating the variability of the estimates across different choices of the synthetic cohort. Although this approach could be used to evaluate the impacts of land rights policy interventions at an aggregate level, we are not aware of any study having implemented it.

3.3 Different policy interventions and evaluation approaches

Table 4 below, based on the arguments discussed above for determining an adequate evaluation method, succinctly presents some possible approaches that can be pursued for evaluating different policy interventions. In general, the choice of evaluation method will be very case specific, and there are no general rules regarding which method to use for which type of intervention. The table is therefore merely suggestive of possible approaches for certain cases. RCTs, RD design or (if the former are not feasible) DiD evaluations are most suitable for

interventions that affect land rights at the individual or micro level. Encouragement design (through complementary information campaigns) can be considered for some components that might in theory affect everybody (such as changes in land legislation). And more aggregate methods such as synthetic cohorts or qualitative evaluations need to be used for interventions that affect everyone at once.

As the most rigorous methods are more easily applicable to interventions that can have variation at the micro level (e.g. mapping or registration), more empirical evidence also exists on those types of interventions. This clearly indicates the need to attempt to include the other more macro-level components in future evaluations. One possibility in programs that include both micro and macro components, is to systematically explore the synergies between both levels in the evaluation. For example, by analyzing heterogeneity of a mapping intervention in some contexts where local institutions have been strengthened and other contexts were they have not, or in some groups who had their rights affected by a given legal or policy change and others who had not, it may be possible to shed quantitative light on the possible complementarities between those interventions. Good qualitative work can complement such evaluations, and might be the most suited to shed light for instance on the results from national level institutional strengthening and policy changes.[12]

3.4 Measuring and controlling for spillover effects

Measuring and controlling for spillover effects is an important part of evaluating the impact of any land administration intervention, and the data collection strategy must account for this. One way of doing so is to consider heterogeneity amongst the control group, particularly spatial heterogeneity if it is believed that those closer to the targeted areas are more likely to feel spillover effects. Spillover effects can also be minimized by choosing control groups that are less likely to be contaminated, such as those further away from the treatment group or under the control of different local governments. But this can conflict with the need to control for unobserved heterogeneity; in practice, a smart design that combines two control groups, with and without exposure to externalities, can allow both identify an intervention impacts and detect externalities. Anticipation effects may also be measured by having one control group that knows it will be treated in future, and another one that does not (see, for example, the Haiti case-study

[12] Guidelines for such qualitative work are beyond the scope of this paper.

Table 3: Mapping of projects to possible quantitative impact evaluation strategies

	Randomized Controlled Trial (RCT)	**Regression Discontinuity (RD) design**	**Difference-in-difference (DiD)**	**Encouragement design (randomized)**	**Synthetic cohort or qualitative evaluations**
Land legislation change		Possibly if beneficiaries identified on transparent rules with clear cutoffs		Possibly if legal change can affect individuals outcomes, by randomizing information about the legislative change	Synthetic cohort approach possible if the intervention is at the local level and some aggregate data is available
Institutional strengthening (including IT)	Only if strengthening of local administrations and number of those high enough	No, unless the institutional change benefits specific groups identified on transparent rules with clear cutoffs			As above
Introduction of digital information technology (IT) for managing cadastral and registry information		As above			As above
Systematic regularization	Yes taking advantage of sequential phase-in, at the locality/commune level. Need to adapt design to possible externalities (e.g. with a 2 stages randomization or a complementary non experimental comparison group).	Possibly if allocation rules or borders of intervention areas, and spatial-continuity assumption acceptable	In case RCT or RD design approaches are not feasible, taking advantage of sequential phase-in, and when common trends assumption acceptable and data collection not too costly. Can provide a complementary non experimental comparison group to monitor externalities.		No, unless all other approaches are not feasible

Cadastral survey and mapping	As above	As above	As above	Only if surveying does not cover all parcels and take-up can be influenced	As above
Land-titling	Yes taking advantage of sequential phase-in, possibly at the individual level. Need to adapt design to possible externalities.	As above	As above	Possibly by randomly subsidizing individuals' take-up if titles have some costs	As above
Registration	As above	As above	As above	As above if registration has some cost	As above

below). Exploiting such heterogeneity is most likely to be possible when spillover effects are foreseen prior to the design of the evaluation. Indeed, the optimal way to ensure such heterogeneity may be through using a two-stage randomized control trial. In this case, the evaluators first randomly select geographical areas to be benefited by the intervention (in which the spillovers are more likely to take place) and then within these areas they pick individuals or smaller areas that will benefit directly from the program. In this way, it can be assured that there will be potentials for control observations both within and outside of the areas in which there may be spillover effects.

3.5 Complementarity with other interventions

The methods discussed above in general aim, at best, at obtaining a credible average treatment (in the case of RCTs, DiD approaches, or regression based estimates) – or in some cases only give a local treatment effect (in case of IVs, and RDD). Yet a natural question is often to what extent impacts may differ depending on the presence of other interventions, or may vary by types of farmers. Such heterogeneity of effects is not only relevant for policy makers, who might want to know which group in the population benefitted most, but also for researchers, since it may shed light on the plausible mechanisms underlying the effects (see section 2). Heterogeneity analysis can be done by estimating impacts for specific subgroups (e.g. large versus small

farmers). Yet the concerns about unconfoundness underlying the previous discussion are equally relevant for heterogeneity analysis.

Specifically, one of the challenges with trying to analyze heterogeneity is to make sure that the dimension of heterogeneity analyzed is effectively isolating the intended mechanism, and does not pick up heterogeneity between an unobserved but correlated dimension. Therefore, in order to rigorously explore heterogeneity, potentially interesting variables should be identified prior to sampling. By stratifying the sample on those variables, one can ensure balance and enough variation in the particular dimension of heterogeneity of interest: e.g. if we are interested in the differential impact of titling on female owners, the sampling frame should purposely include enough female owners. By defining the dimensions of heterogeneity ex-ante, one can clearly signal the *a priori* hypotheses one sets out to test in the impact evaluation. The need to have enough variation in the variables of interest also implies that sample size might need to be expanded, in addition to affecting the sampling procedures themselves.

When heterogeneity of impacts is expected to depend on interactions with other interventions, impact evaluations can go one step further and specifically analyze the complementarity between interventions, by defining strategies to identify the causal impact of each of the interventions and of the interactions. For instance, if the impact of titling on credit uptake is expected to depend on the availability of credit, a simultaneous evaluation of a titling and a credit intervention would be needed, designed in a way that allows separating their effects and consider their complementarity. Yet, such evaluations so far have been very rare. In non-experimental evaluations, the reason for that is straightforward. As it is typically hard enough to have a credible identification strategy for one intervention, having credible identification of two unrelated interventions with the same data in the same setting is often close to impossible.

However, when evaluations are designed prospectively, and especially when randomized allocation is an option, two interventions could be randomized orthogonally on each other. This then will allow not only an estimate of the impact of each of the interventions separately, but also – if designed with sufficient statistical power – to analyze the complementarities between the two. We discuss an example of such an approach in section 4.

3.6 Data Collection

3.6.1 Baseline and follow-ups (long versus short term impacts)

Data requirements for evaluations of land administration interventions clearly depend on the specific intervention examined, on the impact evaluation method chosen and on whether or not the evaluation is prospective or retrospective.[13] While aggregate and more qualitative data can be used for interventions that affect everyone in broad areas, such as legal and policy changes or institutional or technological strengthening, interventions that include components changing rights at the micro level will require data at the individual or micro level.

For prospective designs (i.e. evaluations planned prior to, and in coordination with program implementation), different evaluation methods call for different types of data. For any non-experimental evaluation collecting detailed baseline is absolutely key, as they can serve to help establish the counterfactual. Indeed, multiple rounds of baseline data (or baseline data with additional recall data) can be helpful as it can help to assess the confoundedness or common trend assumption, or, in the case of matching, to obtain better matches. For experimental evaluations with limited number of units over which one randomizes (clusters), baseline data is also very important in order to verify the baseline balance.[14].

The number and the timing of the follow-up surveys is also a nontrivial question. Many of the expected impacts of land administration programs might take a long time to materialize, implying that the full impacts can only be measured after a long period of time. For example, if increased security increases investments in tree planting, the full potential impact in increased agricultural productivity or welfare from the potential reduction in soil erosion would only be observable after many years. Similarly, increases in transfer rights can help guarantee in the long run that land stays allocated to the most efficient user, but may not lead to substantial short-term changes in land allocation or related efficiency. Hence optimally, impacts could be measured many years after the intervention.

Yet this is not necessarily straightforward for a number of reasons. First, in case of randomized phase-in with longer-term impacts one should try to postpone the phase-in of the

[13] If one sets out to evaluate a policy or program that is already implemented (hence a retrospective evaluation), the evaluation will necessarily be restricted to the relevant available data. In such case, arguably, it is the availability of appropriate pre-and post-program data, together with the question on whether program implementation allows for a plausible non-experimental identification that determines whether an evaluation can and should be conducted.

[14] When the units of randomization are very large, randomization guarantees balance on observable and unobservable characteristics at baseline. Such large samples, however, are rare.

experimental control group as long as possible. But keeping the control group out of the interventions for many years might not be feasible. Second, funding for impact evaluations is often linked to project cycles, making it harder to look at longer-term impacts after the end of an intervention. Third, as more time goes by, the comparability of treatment and control groups might become smaller as they may increasingly be exposed to different shocks or other interventions.

Yet, given the importance of knowing longer-term outcomes, smart designs that combine a variety of methods should be used in order to address these constraints ex-ante. For example, it may be feasible to combine a short-term randomized phase-in design with a longer-term non-experimental comparison group. The short term RCT can then serve to validate the validity of the non-experimental group in the short-run, and hence increase the credibility of the longer-term evaluation.[15] For longer-term follow-ups, risks of attrition and hence serious selection bias also tend to increase, so that evaluation designs ought to carefully plan (and budget!) for migrant tracking. In short-term evaluations, data can also be collected on behavioral and attitudinal changes, which can point to potential longer-term outcomes (e.g. attitudes towards investments and how it relates to tenure security).

On the other hand, in order to learn about mechanisms, take-up, compliance, and perceptions about the intervention, data collection during or shortly after the intervention is often also key. Waiting too long to collect the follow-up data might lead to a failure of understanding the underlying mechanisms.

3.6.2 Sampling and stratification

Sampling to a very large extent is driven by the chosen impact evaluation, hypotheses to be tested on average effects and heterogeneity, expectations on take-up, contamination and spillovers, and by the related statistical power calculations. We have discussed above how to adapt sampling in the main evaluation approaches. For example, if the evaluation relies on an RDD approach, one would want to drastically oversample observations close to the treatment threshold. On the other hand, when matching is the method of choice, typically a large sample needs to be collected, in order to obtain within that sample enough observations that can be

[15] This is the approach taken in the 10 year after evaluation of a CCT in Nicaragua by Barham, et al (2012). For Progresa/Oportunidades, the experimental evaluation was also complemented with a non-experimental one to look at longer-term effects (see Behrman et al, 2009).

identified as close matches to the treatment. As costs of such an approach might be prohibitively expensive, this alone indeed is often a very good reason not to opt for matching.

Some rules apply to the different evaluation approaches. Thus in all cases, if any prior data is available on some of the key characteristics to help determine the sample, this can help. In absence of such data (which is unfortunately quite common), households listings (i.e. census lists) may need to be made first, to help establish the universe of households from which to sample. This will also be important if one wants to stratify the sample based on key characteristics of heterogeneity.

Considerations regarding spillovers will be a primary determinant of the level at which the treatment and control groups are defined. For example, a randomized evaluation of titling of individual plots that is expected to increase activities in the land market among community members, should be randomized at the community level in order to account for such spillovers. Alternatively, one could randomly vary the intensity of treatment within communities, in order to obtain an estimate of the spillover effect. In both cases, the correlation between observations from the same community (cluster) will increase the number of observations needed. And sampling would need to capture both sufficient numbers of direct beneficiaries and of other indirect beneficiaries (see Angelucci and Di Maro, 2010 for further guidelines on how to account for and measure spillover effects in designing evaluations).

Statistical power calculations will need to be conducted for each of the main outcomes of interest. The logic behind these power calculations is similar for all types of evaluations, whether they are related to land administration or not (see Duflo, Glennerster and Kremer, 2008 and references therein). In short, power calculations require establishing the minimum detectable effect (MDE) the evaluation aims to identify. They also require estimates regarding the existing mean and variation in the outcomes of interest, the intra-cluster correlations, and the inter-temporal correlation, all of which imply that prior household survey data from the region of interest can be of great value to determine the desired sample size. For power calculations to be useful, they also need to include realistic estimates of take-up rates, compliance with the evaluation design, possible spillover effects that might decrease differences between treatment and control group, and attrition rates. If the evaluation aims at shedding light on possible heterogeneity along a key dimension, power calculations will need to be conducted for each of

the subgroups (as defined by those dimensions) and on being able to identify (with enough statistical power) differences between them.

3.6.3 Outcomes and questionnaire design

For data collection standard practice is to administer household or individual-level questionnaires. These questionnaires need to be carefully designed to assure that not only intermediate and final outcome indicators (and all the variable to calculate those) are collected, but also all variables that are required to do the analysis (such as variables of baseline characteristics) that one might want to control for, information on take-up, etc. Every question in a questionnaire should be justified as an indicator, needed to create an indicator, or variable for the analysis and this should be noted.

The insights from section 2 on different final and intermediate outcomes can help determine which variables to include in survey instruments. It can also be useful to collect information on outcomes that one does not expect to be affected, as they can serve as useful placebo test and validation of the hypothesized causal chain. In addition, to establish the causal chain, it is key to collect 1) information about mechanisms, such as measurements about perceptions of security, conflicts and the possibility of conflicts (latent conflicts), and measurements of effective control (decision making power); and 2) information about implementation, including compliance, possible contamination, timing and delays, perceptions of fairness, etc.

Clearly, in all cases it is important to assure that outcomes related to the policy objectives of the intervention can be measured adequately. This can be more challenging for the broader final outcomes listed above. Very often, secondary sources might be needed for this, such as tax records, satellite data on environmental outcomes or infrastructure, or local election results. The advantage of such additional data sources is that they also offer an independent and non-self reported data source, and sometimes (e.g. in case of satellite data) are in fact observed outcomes, addressing potential concerns related to bias in survey responses.

In order to be able to use such data, care needs to be given prior to data collection on how such information could be linked with the household-level data (e.g. through GPS coordinates, certain geographic identifiers or possible personal identification information). Where more aggregate data is not available, potentially data can be collected through community-level

surveys. This can help to collect information on infrastructure, public goods management and land planning/local development, on conflicts and their resolution, and on prices.

That said, in some cases the level of randomization (or other exogenous variation) might be too small for looking at more aggregate outcomes (e.g. if randomization is at the village level, while election results are only available at the municipal level, impacts on local election results cannot be established). A different approach is then to carefully think about variables in the household (or community) survey that proxy for changes in behavior or perceptions linked to these broader outcomes (such as attitudes and actions regarding soil conservation, or vote intentions). While such reported self-attitudes might be less reliable than observed final outcomes from other sources, they can nevertheless serve as a useful second-best.

In cases where food security is a particular policy objective, there can be challenges at the micro and more aggregate level. At the micro level, careful attention is needed on what outcomes to measure in order to shed light on household food security. One probably does not want only a snapshot of recent food consumption, and additional indicators such as anthropometric measurements might be needed. At the more macro level, the appropriate level on which to expect variations on food security likely depends on the existing integration of markets (e.g. food prices and availability at the village level can be useful if markets are not fully integrated, so that increased production in one village does not automatically get translated in prices and availability of food in other villages).

Finally, some outcomes of interest, such as investment, are quite broad and can be measured in many different ways. Indeed, for most outcomes one can think of various proxy measures. In those cases, evaluations might tend to "cherry-pick" and only show (ex-post) the selection of proxy variables that shows significant impacts. Clearly such practice should be avoided, as in any large set of outcomes there will always be a few that end up being significant. Therefore clear ex-ante protocols and research plans are important, and the causal chain should help predict on which outcomes to expect which type of results. In addition, when many proxies for similar outcomes are collected, corrections for families of outcomes can be made, following Kling, Liebman, and Katz (2004), or alternatively standard errors of individual outcomes should be corrected for multiple hypotheses testing (e.g. using Bonferroni corrections).

Table 4 Examples for Impact/Results Indicators

Indicator	Formula / Definition	Frequency of Measurement	Source
Household welfare	Total expenditures per capita (including home production and bought food); measured at household level	At baseline, first follow-up survey and second follow-up survey	Household survey
Total household income	Total income per capita; measured at household level	At baseline, first follow-up survey and second follow-up survey	Household survey
Agricultural productivity	Total agricultural revenue /hectare (measured at household level)	At baseline, first follow-up survey and second follow-up survey	Household survey
Agricultural investments	Index of agricultural investments (based on reported investments in fixed (e.g. irrigation, terracing, wells, fences) and movable farm assets (e.g. pumps, machinery and animal traction).	At baseline, first follow-up survey and second follow-up survey	Household survey
Natural resource management	Probability and intensity of plot-level long-term productivity-enhancing investments (such as terracing, reforestation, irrigation)	At baseline, first follow-up survey and second follow-up survey[16]	Household survey
Land value	Estimated land value (Prices by hectare) and prices of recent land transactions (including sales, rents, lends, and sharecropping arrangements).	At baseline, first follow-up survey and second follow-up survey	Household survey

[16] The possibility to use geo-referenced aerial photography and/or satellite images to obtain a second measure for environmental impact will be explored (e.g. the quantifiable measure of forest cover).

Which indicators to use for all of these outcomes can be a non-trivial question. Outcomes such as agricultural productivity can be measured in many different ways, and several considerations should be taken into account when determining which indicators to include in data collection instruments. Most importantly, a good understanding of the local context will be needed, and careful piloting of the survey instruments will be key to determine the best way of defining indicators that are good proxies of the outcomes of interest, and that can be measured with sufficient precision. In addition, interest of policy makers in specific indicators, and comparability with other studies should also be carefully considered. Table 4 below provides some examples of possible indicators for various types of outcomes (see appendix 3 for more details).

In addition to outcome variables and variables that might shed light on mechanisms underlying impacts, additional variables that are important to include are those that allow to control for the baseline situation, and if possible, even for baseline trends. One interested in heterogeneity (e.g. by poverty or asset level) indicators of those dimensions need to be included. And in case of non-experimental evaluations, all observable characteristics that might be related to selection into treatment should also be collected. This can hence include a wide set of variables covering a variety of domains such as household composition, past experiences, asset holdings, welfare indicators, shocks, participation in other programs etc. The specific choice will depend on the impact evaluation strategy chosen, the program to be evaluated and the context.

3.6.4 Importance of monitoring and administrative data

Key for any impact evaluation is to be able to document what exactly happened in terms of the implementation of the interventions on the ground. Understanding take-up and possible contamination is not only important to shed light on the magnitudes of impacts and possible biases, but it might also be key to explain potential non-results. For instance, if a large share of the control group in fact receives treatment, or if a substantial share of the treatment group does not, finding non-significant differences between the groups does not necessarily mean that there is no impact of the intervention. Instead, it only means that the non-compliance with the experimental design does not allow one to measure it. On the other hand, if take-up information is available, it can be accounted for in the impact estimates (see above on experimental IV).

While this may seems obvious, it is a point that is often wanting, as good monitoring data that allows one to document the implementation is often missing, incomplete or possibly unreliable. Here ex-ante planning can also be helpful, and impact evaluation teams can try to work with monitoring teams so that the monitoring data is collected in a systematic fashion and in a way that allows it to be linked with the survey data. Such coordination can be a win-win, as for the monitoring team this also provides an opportunity to complement the administrative data with the more detail that micro data household surveys can provide.

3.6.5 Complementary qualitative data

To further understand what happens during program implementation, and how these changes are being perceived by the intended beneficiaries and control groups, qualitative data can greatly complement quantitative evaluations. By focusing on the intermediate steps in the causal chain that might be hard to quantify, it can help to understand *how* the intervention has led to the final outcomes. Qualitative data can further reveal potential unintended consequences (either positive or negative), and as such can help inform adjustment to the survey instruments prior to follow-up data collection. As such, to make the qualitative data collection the most useful, close involvement of the researchers involved with the quantitative evaluation in this part of the research is important.

4. Case study: unbundling rights and identifying complementarities in Haiti.

To illustrate a number of the arguments made above, this section gives an overview of the planned impact evaluation of a land clarification and formalization intervention in Haiti. In addition to isolating the separate impacts of increasing security, individualization and increased transferability, it also aims at analyzing complementarities with other agricultural productive interventions. In this section, we review some of the most pertinent aspects, while the details of the impact evaluation plans are in Appendix 3.

4.1 Evaluation questions on unbundling of rights

First, the overall evaluation question is the impact of increasing tenure security on agricultural investment, productivity, household welfare and natural resource management in selected regions in Haiti. Second, given the pilot nature of the project, the evaluation aims at answering

which mechanism to increase tenure security, and consequently, which component of the overall program design, contributes more to impacts on investments, productivity, welfare and natural resource management, with special attention to the role of clarification of rights, and formalization of rights. Third, what are the additional gains from individualization of rights of family owned or state owned land? And fourth, to what extent impacts of these land rights interventions depend on the presence of complementary productive interventions?

In Haiti, there is very little experience with land titling, and given the presence of many other market failures, it is even qualitatively very hard to predict overall impact of such an intervention. Not surprisingly then there are many open questions on the optimal design of such an intervention, which the pilot and the evaluation are designed to address. Some of the theoretical arguments have some traction in the Haitian context. In particular, land tenure insecurity seems prevalent and concerns both occupants of private and state lands; in this context, some might be discouraged to invest in the land they exploit as the fruits of their investments might be seized by others. Besides, there is anecdotal evidence that the costs of land transactions, at least formal ones, are high, and that this discourages many to fully formalize their rights. Classic theory also emphasizes that improved transfer rights might make it easier for individuals to rent or sell their lands and to others to access some, which could produce both increases in land value and a more efficient distribution of land. Transfer rights in Haiti seem in particular limited for those with land in family indivision and for farmers who rent state land. At the same time, there are some fears that in rapidly urbanizing rural areas, clarification and/or formalization might possible contribute to land speculation, with possible negative consequences for productivity. Finally, others believe that secure property rights are a key element in the fight against erosion, as they provide owners with long-term incentives to protect the top soil.

Titling programs in other countries often take a comprehensive approach combining clarification of property rights, with a complete regularization and formalization of titles. In contrast, the pilot project and evaluation in Haiti was designed to learn which components of such approaches are more instrumental in increasing productivity and welfare. This question is particularly pertinent for Haiti. Some argue that the most important impacts on investments and natural resource management can be obtained through a clarification and documentation of existing rights, without addressing potential land conflicts directly, and without a subsequent formalization process. Such belief is partly based on existing practices in the land markets where

transactions are often not formally registered, but surveying is deemed necessary. Others believe that full formalization is necessary to maximize impacts and that the current practices lead to suboptimal transactions. Yet a trade-off might also exist as formalization of rights, by increasing the stakes, might potentially lead to an escalation of dormant conflicts. This then might have even a perverse effect if it increases overall perceptions of property rights security. In sum, the complexity related to both formal and informal dimensions of property rights, calls for a careful evaluation of costs and benefits related to both approaches. *A priori*, it is unclear whether clarification of rights on its own leads to substantial benefits, or whether the additional benefits of formalization might be higher or lower than those of clarification. This evaluation will shed light on these uncertainties, with the objective to inform future policies.

In Haiti, as in many other developing countries, family land indivision adds to the complexity of property rights. In all countries where by law all heirs are eligible for equal shares of property, often land does not get formally divided, but stays in the name of the entire inheritance (with an existing title remaining in name of the deceased former owner). As a result, none of the heirs obtains a complete property right to their share, as any transaction on part of the parcel needs approval of all the different heirs. The increased transaction costs can potentially lead to substantial inefficiencies in land allocation. In Haiti, some believe family indivisions are mostly due to high surveying and notary costs related to the subdivision, and that, consequently, clarification and formalization of the individual rights on such land could be an important benefit from the titling process. Yet, others argue that land indivision is instead motivated by a mutual insurance rational: family land typically gets used by those family members who stay in the rural areas, without contributions to co-owners that have moved to the cities or abroad. However, for those migrants, the family land serves as an important insurance mechanism, as in times of hardship, they can reclaim their rights by moving back to the family land. Consistent with such a mechanism, providing an opportunity to formalize subdivisions of such family land, might then not lead to more subdivisions or changes in use. Some also fear that providing this option might in fact upset an existing second-best equilibrium, and lead to family conflict over subdivision and/or loss of access to land for the current users. Given these questions it is extremely important to understand the specific impacts for indivision land that could result from the process of clarification of rights and formalization. The evaluation will address this question specifically,

and as such provide some of the first rigorous empirical evidence on this important property rights issue.

Another open question concerns state lands. In Haiti, many farmers exploit state land, which was generally initially rented by ancestors and for which the lease was thereafter transmitted to children and grandchildren. The actual and perceived insecurity of the rights to this land is difficult to comprehend, so that it remains debatable whether state farmers are discouraged to make long-run productive investments on this land. However, if some informal transactions on state land seem to some extent and in some regions possible, these lands often remain in the family, and transfer rights seem limited. With inheritances, farmers apparently often end up cultivating several small and distant parcels.[17] Although this remains to be documented, the rights on state land could thus produce a distribution of land that has allocation efficiency costs by limiting the scope for intensification of agricultural and environmental investments. By comparing impacts of clarification of state land co-owned land with impacts on other land, the evaluation will be also be able to shed light on whether and how individualization of rights adds to the impacts of increased security.

4.2 Evaluation questions on complementarity with agricultural productivity interventions.

Finally, the evaluation offers an opportunity to learn about the complementarities between land tenure security and interventions in the field of agricultural extension or environmental resource management. The IDB and other donors (GAFSP, WB) are planning to implement other programs in the same pilot areas. A first program will support the transfer of technologies, including improved seeds and fertilizers, improved pastures and agro-forestry inputs, through targeted smart subsidies to farmers in the Northern area. Beneficiary farmers can decide to some extent their use of the subsidies. A related project seeks to support private extension and innovation services, by providing demand-driven financial support for service providers. It also envisions strengthening an existing public network of local agents trained at identifying pests and providing vaccination services. In addition, another smart subsidy project will support the management of environmental resources in three watersheds, which are also in the pilot areas of the land tenure program. This program seeks to mitigate environmental risks by financing protection infrastructures, supporting the development of sustainable agriculture, and reinforcing

[17] There might be other motives for cultivating distant parcels such as risk diversification and smoothing of labor requirements.

related institutions. As those programs will be ongoing during the implementation of the land tenure interventions in some areas, the interactions between the clarification and formalization of rights and the agricultural investments interventions could provide additional evidence on the hypothesis that tenure security strengthens incentives for investments and make them profitable and sustainable in the long-run. By coordinating the evaluation designs, it could be possible to provide rigorous evidence on both the individual contributions and complementarities between interventions that will be interest for the design of rural development policies in Haiti.

4.3 Evaluation methodology

The proposed evaluation methodology for the land administration intervention is based on a randomized roll-out of the different components of the intervention. Given that the different activities cannot occur all at the same time because of practical feasibility, randomized phase-in is optimal both for a fair allocation and for the impact evaluation.

In particular, the following method is proposed: in a first phase, half of all localities in each of the targeted communes (randomly selected through a public lottery in the presence of comun leaders and based on pairwise matching) will go through the clarification of rights phase (including land surveying). The first follow-up data collection that will follow this phase will allow measuring the short-term impact of the clarification of rights. While certain impacts possibly might only fully materialize after a few years, investment incentives, changes in agricultural practices, and in perceptions about tenure security are expected to occur in the short run. As such the randomized comparison of the treatment and control group will allow for a clean measure of the impact of the intervention. In addition, likely delays in the projects implementation, may allow observing experimental impacts over a somewhat longer time period.

In the second phase, while the clarification of rights is being rolled out to the second half of localities, one comun will be selected for a small-scale formalization pilot. This pilot will consist in having the landowners of all the private parcels located in a spatially restricted area obtain notary acts and register their parcels at the Direction Générale des Impôts (DGI). The area selected will have previously been covered by the surveying and tenure clarification activities of the program. The objective of this pilot is to specifically test the additional gains from obtaining formalized rights, in addition the rights clarification. Within the selected comun, around 1000 households with plots that are, according to the clarification stage, eligible for formalization, will

be selected. A second lottery will select 500 of these households, for whom the formalization process will occur first. A baseline and follow-up survey (before phasing in of the 500 control plots) will then allow to measure, using the randomized assignment, the impact of the formalization of right.

To further unpack property rights, for both the clarification and the formalization stage, special care will be given in measuring the potential mechanisms leading to impacts. For example, the land clarification stage could affect perceived tenure security in a number of ways: It provides information about users' existing land rights; it may reduces the cost and/or increases the expected probability of success in invoking central government enforcement in the case of future conflict (i.e. the courts); it can increase coordination of local enforcement mechanisms; it can reduce expectation of future government claims on the land; and it can inform users of other peoples claims to their land (i.e. reassures that there is no conflict). In the data collection, special care will therefore be given to obtain measures for each of those mechanisms (e.g. by asking detailed questions on conflicting claims, on local conflict resolution, and on expectations), in order to shed light on how the potential impacts came about, but also to help understand the potential lack of impacts in certain dimensions.

The two-staged, two-sample, randomized design will be complemented with selection of a non-experimental comparison group in neighbouring communes with similar characteristics, which will allow to measure the longer-term impact of the intervention on final outcomes using a difference-in-difference approach (possibly combined with household-level propensity score matching). While there are important potential drawbacks of the non-experimental design (most notably because the communes were selected purposefully for their location in key watersheds where other IDB projects are being targeted), it offers the advantage of having a longer-term estimate. Efforts will be made to coordinate data collection with impact evaluations of the other planned IDB activities, which could allow establishing the validity of the common trend assumption for years prior to the intervention. To further facilitate this, the baseline will contain a limited amount of recall information. The non-experimental group can also allow quantifying potential short-term spillovers of the intervention on the experimental control group. On the other hand, if no short-term spillovers are apparent, the combination of the experimental with the non-experimental approach can allow tests of the assumptions underlying the non-experimental

selection, and to quantify the possible biases in case the assumption of common-trend appear (ex-post) to be violated. As such, both approaches are complementary to each other.

The above design is also well suited to analyse the complementarities between the different components of the land security intervention and the other agricultural interventions in the same regions. In particular, evaluation of the smart subsidy program is also envisioned through randomized phase-in and the randomization of these will occur orthogonally to the land tenure randomization. This will allow measuring the complementarities between both interventions.

Due to the specific design of each of the interventions, the beneficiaries that are likely to benefit from each of the interventions are not necessarily the same. In particular, households that will be direct beneficiaries of the smart subsidies are those who both express their demand and are deemed eligible. On the other hand, households that will benefit from the land clarification activities are those with parcels located in targeted communes and without conflict. These groups will likely overlap to some extent, but not completely. This offers both possibilities and constraints for the integration of the impact evaluation strategies.

In particular, to evaluate each of the approaches, two different samples will be needed, as they will need to be representative of the specific population targeted for each intervention. The sample for the smart subsidies consists of producers that applied and were deemed eligible for the transfer (some of whom, through the lottery process, will receive the transfers earlier than others). The sample for the land tenure program would contain a random sample of households in each of the localities targeted by the program together with a random sample of households in similar localities in nearby areas. The timing of the data collection for each of these interventions is also likely to differ, as it is closely related to the start and phasing-in of each of the interventions. For example, for subsidies, households would be sampled for baseline data collection after they have been identified through the demand-driven approach, and the timing of the follow-up data collection would have to be just before the phasing in of the experimental control groups of those interventions. For the land tenure sample, the baseline would be collected just prior to the start of the land surveying interventions in the early treatment localities, while the first follow-up will occur just before the start of the land surveying activities in the late treatment localities.

Nevertheless, the proposed experimental designs, based on the two separate samples will allow measuring both complementarities and spillovers, and hence shed light on the synergies

between the interventions. For example, for complementarities, and given that the land clarification activities affects (almost) all parcels in the targeted localities, and given the randomized phase-in of this intervention, the complementarities between smart subsidies and land tenure clarification can be identified as long as smart subsidies are randomized and distributed before the phase-in of the land clarification late treatment group. Benefiting from smart subsidies after strengthening of land rights could lead to higher productivity gains than without land rights e.g. because higher land rights might make other households apply for the subsidies, and/or lead to more sustainable use of the subsidies (e.g. more investment in tree crops with stronger land rights). Vice versa, benefiting from strengthened land rights after receiving smart subsidies might help guarantee that changes in behavior obtained through the subsidies are maintained in the longer run. Longer-term complementarities can be analyzed based on the quasi-experimental land tenure treatment group in a similar fashion.

Using the random sample of households from the land sample (many of whom will not receive vouchers and might not directly benefit from extension), we can also measure spillovers from the presence of beneficiaries of vouchers and extension in the same localities. In particular, if the smart transfers and extension support are randomized, the density of beneficiaries of such interventions within a locality will also be random (conditionally on the number of eligible farmers), allowing for a clean identification of spillovers.

5. Conclusions: Practical insights and take home messages

1) Evaluations can serve not only to evaluate the overall impact, but also to provide information on optimal design, allow comparison of different designs and possible sequencing, and provide the opportunity to test complementarities with other interventions. This makes them more attractive for policy makers, as they can provide practical lessons. It also makes them more attractive for research, as it opens the black box, sheds light on mechanisms of impact, and therefore provides information on generalizability of results.
2) Interventions often change property rights in several ways, and there are several links in the chain between the intervention and outcome. To understand which components of an intervention have the largest effects, and what are the mechanisms at work, it is important to collect data on intermediate outcomes, such as conflicts and perceived security, agricultural practices and other labor activities, investments, financial and land transactions. In addition,

evaluation designs should plan for explaining the non-results as well as the potential positive results. This is particularly relevant for land administration programs, for which the available empirical suggests that expected impacts seem often lower than what would be theoretically expected.

3) Impact evaluations will probably not be able to provide quantitative measures of the impacts of every aspect of land administration projects. This is true because projects often combine components focusing both on macro aspects (e.g. institutional reform/strengthening, legal reform) and on micro aspects (e.g. clarification/formalization of rights to specific plots of land). The macro-aspects, by affecting large populations at once, typically do not allow finding a relevant counterfactual for a quantitative evaluation. The quantitative evaluation design therefore should rather focus on separately identifying the impacts of the projects components that are targeted at specific plots, households or communities.
4) Qualitative data evaluations designed to be complementary to the quantitative evaluations can greatly increase the value of the later, and augment the overall lessons from the evaluation by allowing to answer both how and how much the intervention affected expected outcomes. Given the complexities in the causal chain related to land titling programs (where latent conflicts and perceptions of tenure insecurity can sometimes be hard to capture in quantitative surveys) this can be especially important. Early qualitative evaluations and monitoring can also provide key feedback to the program teams, as such increasing the value of the evaluation to them, and possibly enhancing their continued buy-in into the evaluation.
5) Early planning is key. The design of the evaluation and the design of the land rights intervention itself have to go hand-in-hand in order to optimize learning. It is therefore important to get the buy-in for the intervention of all actors and counterparts, and the intensive interaction of the impact evaluation team with the project development team at early stages can facilitate synergies. The time schedules of the collection of impact evaluation data and of the operations have to be defined precisely in advance and made compatible. For instance, the implementation of a titling program based on randomized phase-in requires that the intervention is phased-in in different areas following rigorously the pre-defined schedule (e.g. with two sweeps), and that the follow-up survey be fielded before control areas are phased-in. Early planning allows one to build potential safeguards for the evaluation into the operational procedures, e.g. by writing in compliance with the

randomization procedures into program manuals and terms of references of implementing agencies.

6) High quality monitoring and administrative data are necessary inputs for good impact evaluations. Key for any impact evaluation is to be able to document what exactly happened in terms of the implementation of the interventions on the ground. Understanding take-up and possible contamination is not only important to shed light on the magnitudes of impacts and possible biases, but it might also be key to explain potential non-results. Impact evaluation teams should try to work with monitoring teams so that administrative and monitoring data is collected in a systematic fashion and in a way that allows it to be linked with the survey data.
7) Planning for impact evaluations can gain a lot from being realistic about methodological, operational and political constraints. Careful consideration of risk and opportunities is needed, possible expected and unexpected outcomes should be considered, and all possible biases should be thought of ex-ante so that methods to test for them and avoid them can be built into the data collection, sampling and overall evaluation design. For example, in a randomized evaluation of a land administration programs that involves surveying of plots, the randomization needs to be set up in way that accounts for the fact that such surveying is expensive, takes time, and needs to be organized in a systematic fashion. Clearly randomization at the plot-level will not be feasible, and the unit of randomization will need to correspond to the work zones that are practical for land surveying purposes. Similarly, given that land rights tend to be a politically very sensitive topic, possible political pressures and concerns that might affect compliance with the evaluation design should be carefully evaluated ex-ante, so that the designs can try to minimize the possibility of such pressures.
8) Smart designs combining several methods might be needed to obtain rigorous estimates of both short and long-term results, and both are key for good evaluations of land titling programs. RCTs bring strong internal validity and are the most rigorous approach; they also allow (notably through orthogonal randomizations) evaluating the separate impacts and complementarities of different interventions, e.g. property rights and agricultural development interventions (such as subsidies, access to credit or extension). RCTs can be complemented with non-experimental control groups for evaluating long-term impacts. This will also allow testing for the presence of possible spillovers and side-effects (e.g. information campaigns) of interventions.

Project leader's list of steps to take/decisions to make (often with evaluation team):

1) Identify a principal investigator to work with the program team as early as possible in the process (and prior to finalizing project design).[18]
2) Develop the causal chain and results matrix. Note that there may be several different causal chains, each of which may be interesting to test.
3) Clearly outline the different components of the intervention, and decide for which components there can be sufficient variation in timing or geographical coverage to warrant an evaluation.
4) Derive hypotheses on how each of these components may affect property rights.
5) Informed by 2-4, define the main evaluation questions, including possible heterogeneity.
6) Consider spillover effects and operational logic to determine the level of aggregation at which to measure impacts.
7) Understand program allocation rules, and the potential to influence them, in a way that allows for the definition of a valid comparison group and identification of program impacts over a sufficiently long period. Based on those insights, decide between an experimental or non-experimental evaluation design.
8) For experimental designs, carefully consider threats to internal validity – such as possible low take-up, operational compliance with experimental design, spillovers, etc… – and design the evaluation addressing those possible threats.
9) For non-experimental designs, carefully consider the key identifying assumptions and collect as much *a priori* background information, and among a sufficiently large sample, to evaluate the plausibility of those assumptions.
10) Line out the key requirements and operational constraints related to the impact evaluation for the program evaluation team, including compliance, timing, etc.
11) Work with program team to set up a good monitoring system and systematic administration data, allowing monitoring data to be matched with evaluation survey data.
12) Define sampling frame, power calculations and related budget for data collection.
13) Define the role, team and methods for complementary qualitative evaluations.
14) Define outcomes and indicators, and carefully pilot survey instruments to adapt to local contexts.
15) Organize collection of baseline data.
16) Monitor compliance with evaluation design. In an experimental design in particular, check that randomized determination of eligibility is rigorously implemented, and that the timeline of incorporation of control groups is respected.
17) Implement qualitative evaluation prior to follow-up data collection, to document processes and identify potential unexpected outcomes.
18) Organize collection of follow-up data.
19) Analyse the data from evaluation surveys, estimate program impacts, and document the evidence on the mechanisms at work, seeking to corroborate the causal chain spelled out in 2.
20) Disseminate the findings of the impact evaluation.

[18] Most of the steps in the list will be taken in collaboration with the evaluators/principal investigators. It might be useful to create two separate checklists of steps-to-take, one for project leaders and one for principal investigators This could help assure that the operational constraints and the restitution tasks associated with the evaluation are taken into account by evaluators, and reciprocally that the evaluation needs are taken into account by project teams.

References

Abadie, Alberto & Diamond, Alexis & Hainmueller, Jens, 2010. "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program," *Journal of the American Statistical Association*, American Statistical Association, vol. 105(490), pages 493-505.

Alchian, Armen A. and Harold Demsetz. 1972. "Production, information costs, and economic organization." *The American Economic Review*, 62(5): 777-95.

Ali, Daniel Ayalew, Klaus Deininger, and Markus P. Goldstein, 2011. Environmental and Gender Impacts of Land Tenure Regularization in Africa: Pilot Evidence from Rwanda. World Bank Policy Research Working Paper Series (August 1, 2011).

Angelucci, Manuela and Vincenzo Di Maro, 2010. "Program Evaluation and Spillover Effects", Inter American Development Bank Impact Evaluation Guidelines. Technical Notes. No. IDB-TN-136.

Ayalew Daniel, Matthew Collin, Klaus Deininger, Stefan Dercon, Justin Sandefur, and Andrew Zeitlin (2011). "Land Titling in Urban Slums: The Demand Curve for Property Rights & the Price of Female Empowerment", unpublished working paper.

Bandiera, Oriana, 2007. Land Tenure, Investment Incentives, and the Choice of Techniques: Evidence from Nicaragua. *World Bank Economic Review*21(3):487-508.

Banerjee, Abhijit and Maitreesh Ghatak, 2004. Eviction threats and investment incentives, *Journal of Development Economics*, Volume 74(2): 469-488,

Barham, Tania, Vincenzo di Maro, Karen Macours and John Maluccio, 2012."Assessing Non-experimental Impact Estimates Using Randomized Phase", mimeo.

Behrman, Jere R., Susan W. Parker, and Petra E. Todd. 2009a. "Medium-Term Impacts of the Oportunidades Conditional Cash Transfer Program on Rural Youth in Mexico," in Stephan Klasen and Felicitas Nowak-Lehmann (eds.), Poverty, Inequality, and Policy in Latin America. Cambridge: MIT Press (pp. 219–270).

Besley, Timothy J. 1995. "Property rights and investment incentives: Theory and evidence from Ghana", *Journal of Political Economy* 103(5):903-937

Besley, Timothy J. and Maitreesh Ghatak. 2009. "Property rights and economic development."

Boucher, Stephen R., Bradford L. Barham, and Michael R. Carter, 2005. The Impact of "Market-Friendly" Reforms on Credit and Land Markets in Honduras and Nicaragua.*World Development* 33(1):107-128.

Buchanan, James M & Yoon, Yong J, 2000. "Symmetric Tragedies: Commons and Anticommons," Journal of Law and Economics, University of Chicago Press, vol. 43(1), pages 1-13.

Carter, Michael, and Pedro Olinto, 2003. "Getting Institutions "Right" for Whom? Credit Constraints and the Impact of Property Rights on the Quantity and Composition of Investment", *American Journal of Agricultural Economics*, 85(1): 173-186.

Castañeda Dower, Paul, and Tobias Pfutze, 2011. Specificity of Control: The Case of Mexico's Ejido Reform. Preliminary version (November 12, 2012).

Castañeda Dower, Paul, and Tobias Pfutze, 2012. Vote Suppression and Insecure Property Rights. Preliminary version (April 15, 2012).

Conning, Jonathan H. and Partha Debb, 2007. "Impact Evaluation for Land Property Rights Reforms", Hunter College WP 419.

Conning, Jonathan H. and James A. Robinson. 2007. "Property rights and the political organization of agriculture." *Journal of Development Economics*, 82(2): 416-47.

de Laiglesia, Juan R., 2004. "Investment and credit effects of land titling and registration:," *Proceedings of the German Development Economics Conference*, Kiel 2005 10, Verein für Socialpolitik, Research Committee Development Economics.

de Meza, David & Gould, J R, 1992. "The Social Efficiency of Private Decisions to Enforce Property Rights," Journal of Political Economy, University of Chicago Press, vol. 100(3), pages 561-80

De Soto, H. 2000. "The mystery of capital: Why capitalism succeeds in the west and fails everywhere else." New York: Basic Books.

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

Deininger, Klaus, and Juan Sebastian Chamorro, 2004. Investment and Equity Effects of Land Regularisation: The Case of Nicaragua. *Agricultural Economics* 30 (2):101-116.

Deininger, Klaus, Gershon Feder, 2001. "Land institutions and land markets," in *Handbook of agricultural economics*, eds: L. Gardner Bruce, and C. Rausser Gordon, 288-331: Elsevier.

De Janvry, Alain, Kyle Emerick, Marco Gonzalez-Navarro, and Elisabeth Sadoulet, 2012. Certified to migrate: Property rights and migration in rural Mexico. Preliminary version (March 10, 2012).

De Janvry, Alain, Marco Gonzalez-Navarro, and Elisabeth Sadoulet, 2011.Why are land reforms granting complete property rights politically risky? Electoral outcomes of Mexico's certification program.Preliminary version (August 2011).

Duflo, Esther, Rachel Glennerster and Michael Kremer. 2008. "Using Randomization in Development Economics Research: A Toolkit." In *Handbook of Development Economics*, Vol 4, ed. T.P. Schulz and J.A. Strauss, 3895-3962, Amsterdam: Elsevier.

Feder, G., & Feeny, D. (1991). Land tenure and property rights: Theory and implications for development policy. *The World Bank Economic Review*, 5(1), 135-153.

Fenske, J. (2011). Land tenure and investment incentives: Evidence from West Africa. *Journal of Development Economics*, 95(2), 137-156. Field, 2007. Entitled to Work: Urban Property Rights and Labor Supply in Peru. *The Quarterly Journal of Economics*122(4):1561-1602.

Field, Alfred J., Erica Field, and Maximo Torero, 2006. Property Rights and Crop Choice in Rural Peru, 1994-2004. MTID Discussion Paper No. 100 (IFPRI).

Fort, Ricardo, Ruerd Ruben, and Javier Escobal, 2006. Spillovers and Externality Effects of Titling on Investments. Wageningen University, The Netherlands & GRADE, Lima, Peru.

Galiani, Sebastian, and Ernesto Schargrodsky, 2010. Property Rights for the Poor: Effects of Land Titling. Business School Working Paper, Universidad Torcuato Di Tella, Buenos Aires.

Goldstein, Markus and Christopher Udry. 2008. "The profits of power: Land rights and agricultural investment in ghana." *Journal of Political Economy*, 116(6): 981-1022.

Gonzalez, Mario, Leonardo, Corral, Eric Simning, and Paul Winters, [no date].Evaluability Annex PROCAR II.

Hardin, G. 1968. "The tragedy of the commons." *New York*.

Imbens, Guido W. & Jeffrey M. Wooldridge, 2009. "Recent Developments in the Econometrics of Program Evaluation," *Journal of Economic Literature*, American Economic Association, vol. 47(1), pages 5-86, March.

Imbens, Guido W. & Lemieux, Thomas, 2008."Regression discontinuity designs: A guide to practice," *Journal of Econometrics*, Elsevier, vol. 142(2), pages 615-635, February.

Kling, Jeffrey, Jeffrey Liebman, and Lawrence Katz. 2007. "Experimental Analysis of Neighborhood Effects." *Econometrica* 75(1): 83-119.

Lanjouw, Jean O., and Philip I. Levy, 2002. Untitled: A Study of Formal and Informal Property Rights in Urban Ecuador. *The Economic Journal*, 112(482):986-1019.

Lee, David S. & Thomas Lemieux, 2010."Regression Discontinuity Designs in Economics," *Journal of Economic Literature*, American Economic Association, vol. 48(2), pages 281-355, June.

Macours, Karen, Alain de Janvry, and Elisabeth Sadoulet. "Insecurity of property rights and social matching in the tenancy market." *European Economic Review*, 54(7): 880-99.

Moura, Mauricio, Caio Piza, and Marcos Poplawski-Ribeiro, 2011. The Distributive Effects of Land Title on Labor Supply: Evidence from Brazil. IMF Working Paper (June 2011).

Selod, Harris, Michael O'Sullivan, Markus Goldstein, Kenneth Houngbedji, and Florence Kondylis, 2012. Formalizing Rural Land Rights in West Africa: Results from a Randomized Impact Evaluation in Benin. [Powerpoint presentation]

Sjaastad, Espen and Daniel W. Bromley. 1997. "Indigenous land rights in sub-saharan africa: Appropriation, security and investment demand." *World Development*, 25(4): 549-62.

Vranken, L., Macours, K., Noev, N., & Swinnen, J. (2011). Property rights imperfections and asset allocation: Co-ownership in Bulgaria. *Journal of Comparative Economics*, 39(2), 159-175.

Zegarra, Eduardo, Javier Escobal, Manuel Glave, Ricardo Fort, Ursula Aldana, 2007. "Impactos de la Titulación del PETT en áreas intervenidas de costa y sierra del Perú en el marco del PRTR 2 (periodo 2004-2006), GRADE, report prepared for the IDB.

Appendix 1: Evidence from the empirical literature on land administration interventions[19]

Table A.1 categorizes some of the empirical literature on land rights and land administration interventions by focusing on the evidence on the direct (intermediate) outcomes. In summary, for the four outcomes suggest, the table allows drawing the following conclusions based on a review of evidence available to date.

1. Investment

Some previous evaluations of land administration interventions have found positive impacts on investment. Deininger et al. (2011) and Ali et al. (2011) have found investment alongside households reporting lower perceived risks of expropriation, and the range of investments in these studies suggest it is the greater expected time horizon that is the main channel. Castaneda Dower and Pfutze (2011) on the other hand attribute the greater investment they find to the `reduction in security enhancing actions' channel, since the main investment they find to increase is leaving land fallow which, prior to the intervention, increased the risk of expropriation. As far as we are aware, no study has identified an impact on investment through individualization or facilitation of transfer.

2. Credit

Empirically however, to date there is little evidence of an effect on credit of land administration interventions. Possible explanations include the existence of credit rationing in the countries where impact evaluations have been carried out, or risk aversion on the part of landowners.

3. Transfers of land rights

a. **Sales**

Most empirical studies do not find an impact of land administration interventions on land sales. An exception is Castaneda Dower and Pfutze (2011), who find an increase in sales as a result of the Procede reform in Mexico. They suggest that this is likely to be due to an increase in demand from outsiders as a result of greater security. Lack of evidence for the `facilitation of transfer' channel may reflects the lack of impact evaluations that look specifically at this interventions focusing on this channel.

b. **Rentals**

[19] This appendix and table 2 is not meant to be a comprehensive overview of the literature. Instead, it is based on a set of papers selected for either their focus on Latin America, or for their methodological contribution). It aims at illustrating the type of evidence that exists regarding the different direct outcomes. For a more comprehensive review on the evidence, specifically for SSA, see Fenske (2011).

Several studies have found an increase in land rentals, including Castaneda Dower and Pfutze (2011), Deininger et al (2008), Macours et al (2010) and Alston and Mueller (2012). The increase in security for owners appears to be the mechanism at work in each of these studies.

c. **Non-financial transfers**

Empirically, Ayalew Ali et al. (2011) find an increase in married women's land ownership as a result of the intervention, which was one of the programme's objectives. To our knowledge there is no direct evidence that land has been unintentionally redistributed through land administration programs, though this may be because it is not generally looked for. A piece of indirect evidence is given by Selod et al. (2012), who find that in between knowledge of the intervention and its implementation, land security drops rapidly, suggesting perhaps that many owners fear a resulting redistribution.

4. **Time allocation and migration**

Field (2007) and Moura et al. (2011) find empirical evidence that the land administration interventions they studied did increase labor market supply as a result of increased security. Similarly, Salazar and Quisumbing (2010) find that when women in the Philippines have access to land property rights they are more likely to focus on non agricultural activities. De Janvry et al. (2012) find evidence of increased out-migration as a result of a land administration intervention, but they do not distinguish as to whether this migration is of people who still own the land (and hence a result of increased security) or people who have transferred the land (and hence a result of a the intervention facilitating transfer).

Table A1: Categorization of impacts on intermediate outcomes in empirical work on land administration interventions/land titles in Latin America

	Increased security	Facilitation of transfer	A combination of channels
Invest-ment (in physical capital)	Reduction in security enhancing actions => increase in fallowing and land planted with perennials (Castaneda Dower and Pfutze, 2011) [Non-land investment should not be affected => no increased in current cultivation practices (Castaneda Dower and Pfutze, 2011)] Change in the relative returns to housing investment => improved housing quality (Galiani and Schargrodsky 2001) Spillovers (= learning from others + scale economies for credit institutions) => increased land-attached investments in areas with high titling density (Fort et al. 2006) [Individual titles have no effect on investment (Fort et al. 2006)] Equality of moveable and land-attached investment returns rejected => moving closer to balanced investment portfolio => increase in land-related investment (Deininger and Chamorro 2004)		Increased time horizon => export oriented-crops (Field et al. 2006) [No effect on other agricultural investments (Field et al. 2006)] End of "use it or lose it" rule => increased (Reduced) farmland in high- (low-) productivity areas (de Janvry et al. 2012) Longer contract durations => increased time horizon => Tenants less likely to grow tree crops than owners; (Bandiera 2007) [Moral hazard => type of tenancy contract not correlated with tree cultivation (Bandiera 2007)]
Credit	Increase in mortgaging but only for subsamples (Galiani and Schargrodsky 2010) [Importance of land as insurance? => no increase in access to credit; small increase in mortgaging probability (Galiani and Schargrodsky 2010)] Operational costs of providing loans => Positive relationship between opening of local bank and titling density (Fort et al. 2006) [Investment not working through a "credit-market effect" (Deininger and Chamorro 2004)]		[Possibility to put up land as collateral (Field et al. 2006)] *[Non-price rationing => No impact on formal credit market participation, except for land-rich households; impact even negative for poorest 40% in Honduras (Boucher et al. 2005)]*
Transfer of effective rights	Increased leasing out (Castaneda Dower and Pfutze, 2011 ; Deininger et al, 2008 ; Macours et al 2010; Alston Mueller 2012) [10-year restriction on rentals => no increase in rentals (Galiani and Schargrodsky 2001)] Non-transferable rights creating transfer uncertainty => Titling should increase both sales and rentals markets (Lanjouw and Levy 2002)	[Expiration of the 10-year restriction on rentals => No Increase in rentals (Galiani and Schargrodsky 2001)]	Ability to buy and sell land => enhanced price responsiveness => Export oriented-crops (Field et al. 2006) [Land transactions (Field et al. 2006)] *[History of land reforms? => Increase in land market participation but very low level Distribution of land operated almost unaffected (Boucher et al. 2005)]*
Migration, time allocation and investment in human capital	Reduction in security enhancing actions => Increase in total household labor supply; reduction in probability to run business from home; reduction in child labor (Field, 2007) Increase in secondary and tertiary education (Galiani and Schargrodsky 2010) *Reduction in security enhancing action => Increase in total household*		End of "use it or lose it" rule => Increased out-migration (de Janvry et al 2012) – could be facilitation of sale or reduction in security enhancing actions

	labor supply, significant only around and below the median (Moura et al., 2011)		
Increased land / dwelling value	High and significant titling premium; but insufficient to cover inheritance and other transaction fees (Galiani and Schargrodsky 2001) Transfer uncertainty => Important titling premium, however dampened by strong non-transferable informal rights (Lanjouw and Levy 2002) Increase in self-assessed land price (Deininger and Chamorro 2004)		[Value of dwelling (Field et al. 2006)]
Income / consumption / expenditure	[No increase in household income, and household head income and employment status (Galiani and Schargrodsky 2001)]		[Household total expenditure (Field et al. 2006)]
Health	Investment in human capital and housing improvements Increase in short-term health status measure (Galiani and Schargrodsky 2001) [Investment in physical capital over human capital? => No increase in long-term health status measure (Galiani and Schargrodsky 2001)]		
Politics and conflict	End of vote suppression through expropriation => Increased total electoral participation and votes for opposition (Castaneda Dower and Pfutze 2012) Electoral gratefulness for incumbent party (Castaneda Dower and Pfutze 2012 ; mechanism unclear)		Investor-class and vested interest theories Increased vote share of pro-market party (de Janvry et al. 2011) [Theory of distributive politics => No "gratefulness effect" benefiting the incumbent (de Janvry et al. 2011)]
Demographics	? Reduction in occurrence of teenage pregnancy (Galiani and Schargrodsky 2001) Reduction in security-enhancing actions + Titling as insurance => Smaller, more often nuclear households; reduction in fertility (Galiani and Schargrodsky 2010)		End of "use it or lose it" rule => Reduction in population in poorer areas (de Janvry et al. 2012)
Gender	[No effect on female labor supply (Field, 2007)] Weak informal rights => Female-headed households cannot rent out without ownership title but can easily sell, larger effect of titling on property value (Lanjouw and Levy 2002)		

Notes: Papers are classified according to the aspect of land right considered and the main outcomes the paper analyses. The potential links between mechanisms and measurable impacts that have been identified by theory. Impacts that are `positive' (i.e. roughly equivalent to welfare enhancing) are displayed in green, and impacts that are negative are displayed in red. Black characters in square brackets mean insignificant results, and simple black characters without brackets indicate results that have indeterminate or unclear effects on welfare).

Appendix 2: Impact evaluation plan Haiti

1. Main Evaluation Question(s)

The evaluation is designed to answer a number of questions. First, the overall evaluation question is the impact of increasing tenure security on agricultural investment, productivity, household welfare and natural resource management in selected regions in Haiti. Second, given the pilot nature of the project, the evaluation aims at answering which mechanism to increase tenure security, and consequently, which component of the overall program design, contributes more to impacts on investments, productivity, welfare and natural resource management, with special attention to the role of clarification of rights, and formalization of rights. Third, what are the intermediate behavioral mechanisms that help understand the impacts on final outcomes, such as changes in land management and agricultural investments, changes in agricultural systems, crop and cultivating choices, reallocation of land through the land markets, and access to credit, perceptions of tenure security and changes in family land indivision? And fourth, what are the other expected and unexpected economic and social changes due to different program components?

One of the most common interventions to increase tenure security are land titling programs that formalize property rights. While the IDB and other donors have invested in titling programs in many countries in the region, rigorous quantitative evidence on the impact of such interventions is rare, in particular for rural areas. There is, to the best of our knowledge, no evidence of such programs based on randomized allocations. And existing evidence from non-experimental studies suggests rather mixed results. For example, Carter and Olinto (2003) and Boucher, Barham, and Carter (2005) find that the effects of formal titling might be limited in the presence of other market imperfections. Macours, de Janvry, and Sadoulet (2011) find that informal property rights and conflict affect land allocation, while not finding an effect of formal titles. Field, field and Torrero (2006) find that formal property rights in Peru increase production of export-oriented crops, but does not lead to significant impacts on household welfare or on the market value of land.

The lack of consistent and rigorous empirical evidence stands in contrast with strong predictions of the expected benefits of those programs from neoclassical economic theory, which suggests in particular potential gains on investment, credit access and allocative efficiency of the

land markets. In Haiti, there is in particular very little experience with this type of intervention, and given the presence of many other market failures, it is hence even qualitatively very hard to predict overall impact of such an intervention. That said, some of the theoretical arguments have some traction in the Haitian context. In particular, land tenure insecurity seems prevalent and concerns both occupants of private and state lands; in this context, some might be discouraged to invest in the land they exploit as the fruits of their investments might be seized by others. Besides, there is anecdotal evidence that the costs of land transactions, at least formal ones, are high, and that this discourages many to fully formalize their rights. Classic theory also emphasizes that improved transfer rights might make it easier for individuals to rent or sell their lands and to others to access some, which could produce both increases in land value and a more efficient distribution of land. Transfer rights in Haiti seem in particular limited for those with land in family indivision and for farmers who rent state land (see below). At the same time, there are some fears that in rapidly urbanizing rural areas, clarification and/or formalization might possible contribute to land speculation, with possible negative consequences for productivity. Finally, others believe that secure property rights are a key element in the fight against erosion, as they provide owners with long-term incentives to protect the top soil.[20]

In addition, while the complexity of property rights security has been discussed in the literature, existing titling programs often take a comprehensive approach combining clarification of property rights, with a complete regularization and formalization of titles. Yet, which components of such approaches are more instrumental in increasing tenure security is often not well understood. And as a consequence the marginal benefits on productivity and welfare of the different components of such a package intervention are also not well understood. This question is particularly pertinent for Haiti. Some argue that the most important impacts on investments and natural resource management can be obtained through a clarification and documentation of existing rights, without addressing potential land conflicts directly, and without a subsequent formalization process. Such believe is partly based on existing practices in the land markets were transactions are often not formally registered, but surveying is deemed necessary. Others believe that full formalization is necessary to maximize impacts and believe that the current practices lead to suboptimal transactions. Yet a trade-off might also exist as formalization of rights, by

[20] However, there is also anecdotal evidence that people might see tree planting as an activity with low levels of investments (both in terms of capital and labor), which hence could be reduced by more secure land tenure rights.

increasing the stakes, might potentially lead to an escalation of dormant conflicts. This then might have even a perverse effect if it increases overall perceptions of property rights security. In sum, the complexity related to both formal and informal dimensions of property rights, calls for a careful evaluation of costs and benefits related to both approaches. A priori, it is unclear whether clarification of rights on it's own lead to substantial benefits, or whether the additional benefits of formalization might be higher or lower than those of clarification. This evaluation will shed light on these uncertainties, with the objective to inform future policies.

In Haiti, as in many other developing countries, family land indivision adds to the complexity of property rights. In all countries where by law all heirs are eligible for equal shares of property, often land does not get formally divided, but stays in the name of the entire inheritance (with an existing title remaining in name of the deceased former owner). As a result, none of the heirs obtains a complete property right to their share, as any transaction on part of the parcel needs approval of all the different heirs. The increased transaction costs can potentially lead to substantial inefficiencies in land allocation (see Vranken et al. 2011). In Haiti, some believe family indivisions are mostly due to high surveying and notary costs related to the subdivision, and that, consequently, clarification and formalization of the individual rights on such land could be an important benefit from the titling process. Yet, others argue that land indivision is instead motivated by a mutual insurance rational: family land typically gets used by those family members who stay in the rural areas, without contributions to co-owners that have moved to the cities or abroad. However, for those migrants, the family land serves as an important insurance mechanism, as in times of hardship, they can reclaim their rights by moving back to the family land. Consistent with such a mechanism, providing an opportunity to formalize subdivisions of such family land, might then not lead to more subdivisions or changes in use. Some also fear that providing this option might in fact upset an existing second-best equilibrium, and lead to family conflict over subdivision and/or loss of access to land for the current users. Given these questions it is extremely important to understand the specific impacts indivision land that could result from the process of clarification of rights and formalization. The evaluation will address this question specifically, and as such provide some of the first rigorous empirical evidence on this important property rights issue.

Another open question concerns state lands. In Haiti, many farmers exploit state land, which was generally initially rented by ancestors and for which the lease was thereafter transmitted to

children and grandchildren. The actual and perceived insecurity on the rights to this land is difficult to apprehend, so that it remains debated whether state farmers are discouraged to make long-run productive investments on this land. However, if some informal transactions on state land seem to some extent and in some regions possible, these lands often remain in the family, and transfer rights seem limited. With inheritances, farmers apparently often end up cultivating several small and distant parcels. Although this remains to be documented, the rights on state land could thus produce a distribution of land that has allocation efficiency costs by limiting the scope for intensification of agricultural and environmental investments.[21]

Finally, the evaluation offers an opportunity to learn about the complementarities between land tenure security and interventions in the field of agricultural extension or environmental resource management. The IDB and other donors (GAFSP, WB) are planning to implement other programs in the same pilot areas. In particular, the HA-L1059 will support the transfer of technologies, including improved seeds and fertilizers, improved pastures and agro-forestry inputs, through targeted subsidies to farmers in the Northern area. Beneficiary farmers can decide to some extent their use of the subsidies. In addition, another IDB project will support the management of environmental resources in three watersheds (Grande Rivière du Nord, Cayes, and Cavaillon), which are also in the pilot areas of the land tenure program. This program seeks to mitigate environmental risks by financing protection infrastructures, supporting the development of sustainable agriculture, and reinforcing related institutions. As those programs will be ongoing during the implementation of the land tenure interventions in some areas, the interactions between the clarification and formalization of rights and the agricultural investments interventions could provide additional evidence on the hypothesis that tenure security strengthens incentives for investments profitable and sustainable in the long-run. More generally, the individual contributions and complementarities between interventions will be of interest for the design of rural development policies in Haïti.

In sum then the expected benefit of conducting this evaluation is to fill important knowledge gaps, both with the eye on a possible expansion of such interventions in Haiti, and elsewhere in Latin America and the Caribbean.

[21] The questions regarding land in indivision and state land call for heterogeneity analysis along these dimensions (see further). In the Haitian context, there are a priori less reasons to believe there would also be important heterogeneity on household's economic status, and hence this is not necessarily a dimension of heterogeneity we aim to focus on.

2. Existing Knowledge (previous evaluations, ex ante economic analysis)

2.1 Insights from prior studies

Given that this is a pilot program for Haiti, there are no obvious relevant prior experiences to learn from or compare to. That said, concerns about tenure security and property rights have been highlighted in different working papers and project documents of various donors, and in the sociological and anthropological literature on peasant farming in Haiti.Existing studies suggest that the pressure on farmland is high in Haiti (Report from French Cooperation, 2010). Land pressure results from the many difficulties the agricultural sector, dominated by very small farms, faces to increase productivity and meet with food security challenges, and from the deterioration of soils with deforestation and erosion. The demand for agricultural lands further stems from the expansion of cities in peri-urban areas. This demand, in spite of the lack of financing from formal institutions, generates a high number of transactions on land in certain rural areas.

However, although a comprehensive legal framework exists, formal procedures for land transactions (which include the survey of plots by surveyors, certification by notaries, and recording by the Ministry of Finance (DGI) services) are costly and rarely completed entirely. Hence many different property titles have come to coexist (e.g. "surveyor" certificates) (Dorner, 1999). Moreover, the institutions in charge of the management of land transactions are reportedly characterized by a significant degree of disorganization (e.g. many titles and transaction records are not centrally kept and often get lost) and dysfunction (e.g. several different titles can exist for the same plots), and the practical management of land transactions has in many cases evolved independently from formal procedures. The literature thus suggests that land tenure management in Haiti remains largely informal. This informality is reinforced by the sharing of lands in "family indivision", at inheritances, which remains done by successors without any formal surveying or recording (Dorner and Oriol 2009). This informal sharing is apparently often felt as definitive although the land remains legally in indivision, and the situation becomes even more complicated as transactions occur on parts of the indivision plots.

The disconnect between the legal framework and land market reality is generally thought to generate a widespread insecurity on land tenure. Indeed, although the recording of courts decisions is lacking, tenure conflicts are reported to represent a large share of the cases to be decided by local courts (Bucazelle, 2004). Tenure insecurity extends to state farmers, who own

leases for lands in the private domain of the state. Although those farmers are reported to perceive a substantial degree of security on this tenure, those leases can theoretically be interrupted by state services, and acquisition by prescription is not an option on state land. Moreover, informal division and transactions have apparently also developed on farmed state land, as the original leases were transmitted to children and great-children.

While no quantitative evidence is available, several studies hypothesize that this generalized insecurity negatively impacts investments in agricultural production and the management of environmental resources (Dorner, 2004; Smucker, White and Bannister, 2002). Consistent with such arguments, Verner (2004) reports a strong correlation between the ownership of formal rights on farm land and agricultural incomes.

2.2 Insights form the ex-ante analysis

More directly related to the program, the ex ante economic analysis of the pilot project itself raises a number of important questions regarding the potential benefits of the increasing land tenure (Couralet, 2011). It suggests that only relatively modest gains on income might result from formalizing land titles, with somewhat stronger gains for land in indivision and state land. The results on investments (tree planting), in line with the literature, also indicates that gains might not come automatic.

Nevertheless, the findings of the ex-ante evaluation come with a number of serious caveats. It is based on a single cross section from 2001 and without plot-level information, and hence the level of analysis, the potential for causal inference, and the relevance for the current (post-earthquake) situation is necessary very limited. Disentangling the impact of different components of a potential tenure program is also not possible with this data, nor is shedding light on different short-term versus longer-term impacts.

Given the many open questions resulting from the ex-ante evaluation, as well as the social impact evaluation, a rigorous impact evaluation seems crucial in order to obtain reliable estimates of the benefits of this type of intervention in Haiti. This impact evaluation will be informed by the important questions and hypotheses raised in the ex-ante evaluation. These are further enriched by the broader literature on the topic referenced above.

3. Key Outcome Indicators

The objectives of the program are to increase land tenure security of rural households and strengthen institutional capacity for a more efficient management of land tenure. To meet these objectives the program would be structured into two components: (1) land tenure clarification and formalization in targeted areas; and (2) institutional strengthening for land administration. The overall goal of the program is to promote and facilitate medium and long-term investments in agriculture and sustainable land and natural resources management in pilot targeted areas.

The quantitative randomized impact evaluation discussed in more detail below aims at measuring the separate impacts of land clarification and formalization on medium and long-term investments in agriculture and sustainable land and natural resource management by focusing in particular on impacts on agricultural investment decisions and the resulting agricultural productivity and household welfare.

Table A.2 gives the main results indicators that will serve as final outcomes in the impact evaluation. Those include household total consumption and income, agricultural productivity, natural resource management, and land value.

To understand the causal chain of events, impacts on perceptions of tenure security, conflicts, and short-term decision making on crop and cultivation choices, reallocation of land through the land markets, changes in land values, and access to credit will all be quantified. For this purpose, and more succinctly, the intermediate outcomes under scrutiny will in particular include:

a) tenure security perceptions: perceived risks that somebody might claim on ones land, or that the government might take the land away from the individual (at the plot level);
b) frequency and intensity of land conflicts: involvement in land conflicts by type;
c) crop choices: annual crop production and planted areas;
d) input use: use and expenditure on fertilizers, pesticides and other chemicals, and other expenditures in land preparation;
e) labor use: amount of unpaid family, amount and cost of paid paid labor (distinguishing day work and permanent labor);
f) farm yields: annual production and sales by crop;
g) permanent crops production: inventory of trees, planted area, annual production and sales by crop for each plot;

Table A.2 Key Impact/Results Indicators

Indicator	Formula / Definition	Frequency of Measurement	Source
Household welfare	Total expenditures per capita (including home production and bought food); measured at household level	At baseline, first follow-up survey and second follow-up survey	Household survey
Total household income	Total income per capita; measured at household level	At baseline, first follow-up survey and second follow-up survey	Household survey
Agricultural productivity	Total agricultural revenue /hectare (measured at household level)	At baseline, first follow-up survey and second follow-up survey	Household survey
Agricultural investments	Index of agricultural investments (based on reported investments in fixed (e.g. irrigation, terracing, wells, fences) and movable farm assets (e.g. pumps, machinery and animal traction).	At baseline, first follow-up survey and second follow-up survey	Household survey
Natural resource management	Probability and intensity of plot-level long-term productivity-enhancing investments (such as terracing, reforestation, irrigation)	At baseline, first follow-up survey and second follow-up survey[22]	Household survey
Land value	Estimated land value (Prices by hectare) and prices of recent land transactions (including sales, rents, lends, and sharecropping arrangements).	At baseline, first follow-up survey and second follow-up survey	Household survey

[22] The possibility to use geo-referenced aerial photography and/or satellite images to obtain a second measure for environmental impact will be explored (e.g. the quantifiable measure of forest cover).

h) animal rearing: ownership of livestock, annual purchases, consumption and sales;
i) access to credit: formal and informal loans (incl. source and amount, interest rate, and requested collateral), credit rationing status, and perceived credit access;
j) land transactions: parcels bought, rented, inherited, sharecropped, borrowed, and given out (incl. contract type, dates, area, prices, and description of counterparts);

More detailed information on the instrument for collecting the data is given below.

This quantitative approach could be complemented by a process evaluation to shed light both on the operational details regarding the implementation of each of the steps of the intervention, and on the institutional component of the intervention. A targeted qualitative evaluation on expected and unexpected impacts will also precede each of the follow-up surveys in order to enrich understanding of mechanisms underlying potential changes in outcome indicators and to inform changes in questionnaire design and as such complement the quantitative approach.

4. Evaluation Methodology

The main evaluation methodology proposed is based on a randomized roll-out of the different components of the intervention. Given that the different activities cannot occur all at the same time because of practical feasibility, randomized phase-in is optimal both for a fair allocation and for the impact evaluation.

In particular, the following method is proposed: in a first phase, half of all localities in each of the targeted communes (randomly selected through a public lottery in the presence of comuns leaders and based on pairwise matching) will go through the clarification of rights phase (including land surveying). The first follow-up data collection that will follow this phase will allow measuring the short-term impact of the clarification of rights. While certain impacts possibly might only fully materialize after a few years, investment incentives, changes in agricultural practices, and in perceptions about tenure security are expected to occur on the short run. As such the randomized comparison of the treatment and control group will allow for a clean measure of the impact of the intervention. In addition, likely delays in the projects implementation, may allow observing experimental impacts over a somewhat longer time period.

In the second phase, while the clarification of rights is being rolled out to the second half of localities, one comun (or a group of secciones comunales) will be selected for a small scale

formalization pilot. This pilot will consist in having the landowners of all the private parcels located in a spatially restricted area obtain notary acts and register their parcels at the DGI. The area selected will have previously been covered by the surveying and tenure clarification activities of the program. The objective of this pilot is to specifically test the additional gains from obtaining formalized rights, in addition the rights clarification. Within the selected comun, around 1000 households with plots that are, according to the clarification stage, eligible for formalization, will be selected. A second lottery will select 500 of these households, for whom the formalization process will occur first. A baseline and follow-up survey (before phasing in of the 500 control plots) will then allow to measure, using the randomized assignment, the impact of the formalization of right.

Finally, it is suggested to complement this two-staged, two-sample, randomized design with selection of a non-experimental comparison group in neighboring communes with similar characteristics, which will allow to measure the longer-term impact of the intervention on final outcomes using a difference-in-difference approach (possibly combined with household-level propensity score matching). While there are important potential drawbacks of the non-experimental design (most notably because the communes were selected purposefully for their location in key watersheds where other IDB projects are being targeted), it offers the advantage of having a longer-term estimate. Efforts will be made to coordinate data collection with impact evaluations of the other planned IDB activities, which could allow establishing the validity of the common trend assumption for years prior to the intervention. To further facilitate this, the baseline will contain a limited amount of recall information. The non-experimental group can also allow quantifying potential short-term spillovers of the intervention on the experimental control group. On the other hand, if no short-term spillovers are apparent, the combination of the experimental with the non-experimental approach can allow to test the assumptions underlying the non-experimental selection, and to quantify the possible biases in case the assumption of common-trend appear (ex-post) to be violated. As such, both approaches are complementary to each other.

The above design is also well suited to analyze the complementarities between the different components of the land security intervention and the other agricultural interventions in the same regions. In particular, the possibility of randomized phase-in of the IDB-funded voucher for technology adoption is being explored. If possible, this randomization will occur orthogonally to

the land tenure randomization, and as such will allow measuring the complementarities between both interventions (see appendix 1).

5. Technical Aspects of Selected Methodology

5.1 Evaluation of impact of clarification of rights

1) The evaluation of the clarification of rights will rely on randomized roll-out of the land surveying activities, with treatment and control groups defined at the level of the locality/habitation. The intervention is planned to take place in 10 communes, belong to 3 different departments and 2 watersheds (the Grande Rivière du Nord watershed in the North and Northeast departments and Cayes/Ravine du Sud watershed in the South department). (Each comun can be subdivided in "sections communales", which is the smallest administrative unit in Haiti. There are between 1 to 8 "sections communales"in the communes that have been selected.) The operational field work for the cadastre, and hence the evaluation, will be organized by locality (or "habitations"). The locality is a geographical sub-unit of the section communal, and corresponds to settlements as established during colonial times.[23]
2) The assignment of early treatment and late treatment groups will happen through a public lottery organized in each region, to which the mayors and elected officials of all "sections communales" will be invited. The randomized allocation will rely on pairwise matching: i.e. before the lottery, localities will be organized in pairs (based on similarities), and from each pair one locality will be randomly selected to be in the early treatment groups, while the other locality will be in the late treatment groups. Given that the estimated number of localities is more than 500 this strategies ex-ante maximizes the probability of having statistically similar treatment and control groups, while the stratification that is implicit in pairwise matching allows investigating geographical heterogeneity ex-post. The later is important given that expected impacts could be different related to topography (e.g. impact on erosion reducing investments (in the mountains) versus land speculation (in the plains)).
3) The non-experimental comparison group will be established in two phases. First, localities in neighboring communes will be selected based on available data from the agricultural census,

[23] While the localities do not correspond to a clear administrative entity, the limits of the localities are known at the local level, and each plot can be clearly assigned to locality. The project envisions the delineation of the localities in preparation of the operational work, which will occur using local knowledge.

population census, and data collected in the framework of other IDB interventions. Second, after the baseline data is collected, household and/or plot-level propensity score matching (based on baseline variables as well as recall information) will be used to refine the comparison group

4) Because some of the program components and effects are likely to affect behaviors of households and farmers whatever the locality they live in, it will be important to distinguish the direct effects of the surveying of the plots owned by individuals in the treatment group from broader program spillover effects. Those spillovers are likely to occur in particular as the public awareness and information campaigns on the importance of secure lands rights and responsibilities of land owners and occupants that precede the pre-cadastral sweep will cover the entire pilot communes, and the clarification or registration of rights process in some localities might affect overall perceptions and anticipation of future tenure security in yet uncovered localities. The suggested evaluation strategy will help to disentangling direct effects from indirect and/or spillover effects. Indeed comparing outcomes of the two experimental groups will allow isolating the impact of the clarification of rights on individuals' parcels in se from the spillover effects, while the non-experimental comparison group will permit estimating the extent of spillovers on the controls, which are of proper interest.
5) However, a much more severe contamination would occur if any of the control localities are included in the first-phase cadastral sweep. The timing of the intervention and evaluation surveys will hence be critical to avoid this, with the first follow-up going on field before the first stage component has begun in the control group, and the contract with the firm in charge of managing the first stage clarification of rights will have to incorporate this sequencing (see more on the timing of the surveys below).

5.2 Heterogeneity and impact of clarification on specific types of plots:

6) *Plots under conflict*: Because some land tenure conflicts will remain unsettled after the clarification of rights first stage of the program, some households living in program localities won't be eligible to, and therefore able to reap any benefits from, the formalization of rights second stage. This does not pose a compliance issue for the second stage, as the 1000 households will be chosen among those eligible. However it will be important to measure a

potential adverse effect of the clarification stage on conflicts. The impact evaluation and the data from the cadastral sweep, which identifies the parcels under conflict, will allow examining these effects of the program on this sub-group of plots by comparing the outcomes of parcels under conflict after (treatment group) and before (control group) the conflict has been detected in the clarification of rights process.

7) *Family land in indivision and state land*: The baseline survey will identify plots that are in family indivision as well as state land in both the treatment and control groups. By considering the impact on this subset of plots, and comparing it with other plots, we can shed light on the potential gains and costs related to clarification of rights on these plots. Based on the available data it is hard to predict the share of plots in indivision or state land. Based on anecdotal evidence, there are strong reasons to believe that the share of land in indivision is substantial but the data from the agricultural census almost certainly provide a very serious underestimate of land in indivision.[24] The same holds for state land. In the baseline survey, careful attention will be giving in piloting and training of field personnel, to guarantee a complete identification of the ownership and tenure status of each plots, which will allow to obtain more reliable estimates. As such, this will be one the contribution of the baseline survey of this evaluation.[25]

5.3 Evaluation of impact formalization of rights (small scale pilot)

8) The approach followed for this pilot will consist in a proof of concept: it will be sought to evaluate the relevance of formalization of land tenure rights, in the context of rural areas of Haïti, by examining the benefits it brings in a favorable and controlled setting. The area for implementing this pilot will thus be chosen as exhibiting conducive conditions, notably in terms of local notaries' skills and willingness to collaborate, and of local DGI services' functioning and resources. In case this pilot and its evaluation reveal large benefits from formalization of tenure rights, these findings will serve as inputs for setting up future policies

[24] E.g. according to the census, 60% of land in the 10 comunes was bought with a title. This certainly must be too high to only include land bought by the current owners, and hence is likely to include many plots in indivision. According to the census, only 1.5 % is state land, but again this seems very low, compared to causal observations in the regions of study.

[25] Even though for state land, the precision of those estimates might be low in case they are geographically concentrated. Overall, given the uncertainties on the share of state and indivision land, we do not perform separate power calculations for heterogeneity. Note however that state land tends to be geographically concentrated, and the pairwise matching of localities should hence help to ensure a balance of the share of land in state land between treatment and control. The power of the heterogeneity analysis will be reconsidered after the baseline survey.

for formalization on a large scale. The implementation of such large-scale policies, which would likely differ from those of the small-scale pilot, would be examined only at that later stage.

9) The geographic area and the number of beneficiaries of this formalization pilot will be restricted, so as to limit the costs and complexity of implementation operations (the high level of notary fees make the scaling-up of such activities costly in the current setting), though sufficiently large to allow for a quantitative evaluation of its impacts. In practice, this component will cover one commune (or a set of adjacent sections communales) in which about a thousand landowners hold incomplete land tenure titles, i.e. title that do not come with notary deeds. All the landowners of the parcels located in that area will benefit from the formalization activities, which will insure equal treatment within the pilot area.
10) The evaluation of this component also relies on randomized phase-in at the household level. Through a second lottery, half of the eligible households (approximately 500) will benefit from formalization as soon as possible following clarification. The other half will receive the same benefits 2 years later. A baseline and follow-up survey (before phasing in of the 500 control households) will then allow to measure, using the randomized assignment, the impact of the formalization of right.
11) The formalization will begin in the third year of the program, once a commune (or a set of adjacent sections communales) will have been entirely covered by the surveying and tenure clarification activities of the program. The parcels eligible for formalization in those areas will be identified, after the surveying activities are completed, as all private parcels that were surveyed and tenure-clarified, but for which landowners do not possess any notary act. Parcels object to a conflict will not have been surveyed and will also be excluded from formalization. Tenure rights will be registered by notaries and DGI for about a thousand parcels, and every landowning household will benefit from formalization for at most one parcel.
12) The proposed impact evaluation will rest on the need to spread the formalization over several years, due to the incapacity of local notaries to handle a large number of land tenure formalization cases in a limited period of time. The sequence for the processing of the cases and the delivery of notary acts for each parcel will be determined through the public lotery. Half of the parcels, that is around 500, will be selected for formalization during the first two

years of that component (i.e. third and fourth years of the program) while the other half will get formalized only after the fourth year. The lotery for determining the sequence will be hold publicly at the council of the commune selected for benefiting this component.[26]

13) Survey data will be collected among the households of the sections communales benefiting from the formalization component before those operations have begun, i.e. during the third year of the program, and at midterm after the deeds have been issued for the first half of the parcels to be formalized, but before the notaries begin to issue any title for the second half of the parcels. This quantitative evaluation will allow measuring the impacts of formalization after two years.

14) This component and its evaluation will serve to inform on the additional benefits from formalizing land tenure, beyond those from land clarification. It should provide indication on the additional land tenure security derived from holding a notary act with respect to only a survey act ("procès verbal d'arpentage"), and its effects on investments and agricultural productivity or on the management of environmental resources (and other results sought by the program).

5.4 Attrition

The relatively high mobility in Haiti, with in particular migrations to urban areas and abroad, could be a source of attrition in the panel of households. This needs to be taken into account for the design of the sample (see below). Although the program might also affect migration decisions, and notably reduce those by increasing current and future land tenure security, - which would be of proper interest -, those effects should be limited in the short-run. Because attrition can lead to selection bias in the experimental and non-experimental estimates, it will be important to reduce it; for instance information will be collected from non movers on migrant relatives, and the survey procedures will seek to locate households who moved in the local areas. More generally, attrition is a central concern for the validity of impact evaluation and it will be important that the procedures of the follow-up surveys make it as small as possible.

5.5 External validity

[26] More details on the number of beneficiaries of this component is given in the section on power calculations below.

The estimates of the impacts of the land tenure regularization program will be obtained for the rural population of 10 communes and three large watersheds of Haiti (Haiti has 33 watersheds and 133 communes). Those represent a sizable share of the national population and territory, and should provide a solid basis for understanding the potential from a scaled-up land tenure security policy. However, the specificities of the selected communes and watersheds might limit the external validity of the results from the impact evaluation. Such external validity concerns are intrinsic to the evaluation of any pilot program. In particular, the pilot communes were selected based on their coverage by other IDB activities, in particular two environmental resource management and agricultural extension programs. The average estimated impacts will thus include some complementarities with the effects of the other programs. However, not all pilot communes are included in the other programs, so the impact evaluation will provide some insight on those complementarities. Moreover the pilot areas cover a variety of geographic and socio-economic conditions, with entire watersheds included from peri-urban areas to highlands, both in the Northern and Southern regions.

The external validity of the small-scale formalization pilot will necessarily be more limited, consistent with its primary objective of a proof-of-concept.

5.6 Power Calculations

The power calculations were performed with data from the Haiti LSMS 2001. This is the most recent household-level data available, and is also the dataset used for the ex-ante evaluation.

5.6.1 Power Calculation for Clarification of rights

The impact evaluation design is based on 3 rounds of the household and plot survey (baseline, mid-term and endline survey), on 3 different groups of households: early treatment group, late treatment group and comparison group.[27] This will allow estimating both the short-term and the longer-term impacts of the clarification of rights. The short-term impacts will be estimated based on the randomized rollout of the intervention. The longer-term impacts will be estimated based on a non-experimental comparison group and a difference-in-difference matching estimator.

[27] The baseline survey will be fielded before the clarification of rights stage, takes place in the pilot areas. The mid-term survey will be fielded after the first stage has been completed in the early-treatment group localities and before it has begun in control group localities, while the end-line survey will be fielded after the clarification stage in the late-treatment group. The mid-term and endline survey will be preceded by a targeted qualitative evaluation.

Statistical power calculations were performed to establish the number of households needed in each group and each survey round. The power calculations focused on the expected increases in rural household income, both because increasing income (through increases in agricultural productivity) is one of the primary objectives of the intervention, and because of data availability.[28] The power calculations start from an expected increase in rural household income of 19%, and were done with power of 0.8 and for 5% significance level.[29]

Estimates of mean standard deviations and intracluster correlation (0.139) of income in the 3 departments covered by the pilot were obtained from the most recent household survey with data availability (the Haiti LSMS 2001). Given that no panel data is available, the expected correlation between baseline and control (.45) was estimated based on a recent panel survey in a poor agricultural region of Nicaragua.[30]

The estimated number of localities/habitations in the 10 communes is 540 (estimates from maps of the Haitian statistical Institute) and the estimated number of parcels per household is 2. The estimated number of parcels per household is based on anecdotal field observations and preliminary data from the recent agricultural census. It is important to point out that some impacts will be measured at the parcel level (e.g. plot-level investments) while others will be estimated at the household level (e.g. total agricultural productivity). Households in the region of study often own parcels that are relatively far away from each other, and might be located in a different localities/habitations. This implies that there will be random variation in the share of a household' parcels treated. At the household level this random share will be used as the intent-to-treat estimate, which allows accounting for the intensity households were treated. A priori, it is hard to estimate the share of households that will have plots with different treatment status. That said, it is clear that this provides additional variation in the independent variable. The total number of households to be sampled is therefore estimated to be half of the total number of observations indicated by the power calculations.

[28] With the available data on agricultural productivity itself notably of lower quality.

[29] 19% is higher than the coefficient estimated for the correlation between titled land and income in the ex-ante evaluation, but in magnitude more similar to coefficients for state land and land in indivision, and potentially could be too optimistic. That said, the standard deviations used in the calculations was large, driving up the sample size. To illustrate this, the sample size calculated is also sufficient to find an impact of less than .2 sd of estimated income in the Haiti LSMS. A change of .2 standard deviations is often used as a useful benchmark impact size to find relatively small impacts.

[30] Given that Nicaragua is the second poorest country in Latin America after Haiti, this seems to be a plausible relevant and available correlation to consider.

Note that there will be some localities in which no households are living anymore. The locality maps of the Statistical institute, which date from 1979-1981 indicate 13 localities without households. This number is expected to be higher now, as households might have moved closer to roads and better infrastructure. This implies that we will not be able to sample households in those localities. The strategy for baseline survey data collection will guarantee that any observations that might be missing from the sample because of these issues will be compensated for by oversampling proportionally in other localities of the same treatment group located in the same section communale. In addition, there are likely urban households with rural property that need to be surveyed. Given that a lot of the rural property is owned or used by households living in the comunne's urban areas, this is necessary to capture an important group of beneficiaries, and estimate the overall impact of the intervention. For the power calculations, we anticipate these two effects to cancel out, and hence use a total number of clusters of 540.

Given the relatively high mobility in Haiti, it is important to plan for the fact that a number of households interviewed at baseline might not be found back to re-interview at follow-up. The power calculation therefore assumes a 10% attrition rate between baseline and each of the follow-up surveys.

Taking all these factors into account, we find that the targeted sample size for each group is 1123 parcels, or hence 562 households at baseline. This implies a total of 1685 households with 3369 parcels to be surveyed between early treatment, late treatment and comparison groups.

These estimates assume 100% compliance for the clarification of rights. This is motivated by the fact that the land clarification activities will systematically cover all plots in each locality. That said, if for a subset of plots no productivity gains are to be expected (e.g. because plots are in conflict and no documents will be issued), sample size would need to be adjusted upward to take this into account.

5.6.2 Formalization pilot

The randomization for the formalization will be at the household level. As for clarification, the power calculations focused on the expected increases in rural household income. As before, the power calculations start from an expected increase in rural household income of an additional 19% (in addition to the 19% obtained from clarification), and were done with power of 0.8 and for 5% significance level. Assuming 100% among eligible households, this leads to an estimate

of 508 households per group, and hence 1016 for treatment and control group combined. Some additional power will be gained from locality level averages obtained from the clarification baseline.

5.7 Data Collection

The main data collection instrument for both the evaluation of clarification and formalization will be a household survey with detailed plot-level information. This will be combined with data on exact location, borders and property rights status that will be directly obtained from the field work related to the clarification of rights (in particular cadastral maps with limits, and information on tenure of each plot).

The household survey will contain a number of modules on the household's agricultural activities and investments.

(i) In particular, a plot-level module will enumerate, locate, and record the identifiers for panels of the agricultural parcels that the household owns and/or uses. It will also collect information on the physical characteristics (incl. area, soil type, slope, cultivable area), tenure status (rights, reported owners and users, information on existing renting contracts (if relevant), conflict status, past conflicts and conflict resolution), physical infrastructure (irrigation, terracing, soil improvement, reforestation, well, pumps, other environmental investments) and use in different seasons, of the plots.

(ii) This module will also collect information on the use of chemical and biological fertilizer and pesticides (type, amount and annual purchase price of fertilizers used, annual purchases of pesticides and other chemicals used), and elicit the permanent crops cultivated on each plot (incl. number of trees, distribution by age, production and sales).

(iii) A crop-level module will capture annual crop production and use of different crops (area planted, total production, revenue, own consumption, sales, gifts, and other uses by season) to obtain an aggregate household-level indicators of agricultural productivity and changes in agricultural systems.

(iv) A household-level agricultural module will also collect information on the stock of agricultural machinery and assets that the household owns (incl. their number, current value, recent purchases), and on household-level agricultural production costs, in particular costs in machinery and animal traction (use, ownership/rent, and cost of

machinery and animals by agricultural task) and annual labor costs (amount of unpaid family labor used, amount and cost of paid day or permanent labor used).

(v) A livestock module will elicit the number of animals currently owned as well as those purchased, sold (with revenus), and consumed in the last 12 months.

(vi) A land markets module will shed light on past and current transactions in the land market, by collecting information on formal and informal sales, rental or gift agreements, counterparts in these contracts, costs involved with different transactions and level of formalization, transaction prices; the parcels involved in those transactions will be identified;

(vii) A module on land security and land conflicts will ask about farmers experience in land conflicts (incl. past and/or current involvement in conflict(s), their description, the involvement of relatives, institutions involved in their resolution/management), in perceptions related to tenure security (perceived risks that somebody might claim on ones land, or that the government might take the land away from the individual).

In addition the household survey will contain:

(i) A detailed consumption module to calculate the consumption aggregate, meant to provide an overall measure of household welfare; this will also include a section on household durables, which will allow capturing nonagricultural investments.

(ii) A nonagricultural income and economic activities module collecting information on both wage income (incl. type of job, work hours, total income) and business income (incl. ownership and type of business, value of fixed assets, time devoted to the business, and obtained revenues and profits). This will allow understanding the impacts and complementarities of the agricultural income with other household income.

(iii) A section on household composition, recent migrations, remittances (incl. the identification and basic characteristics of household members who migrated and value of remittances) and household social networks.

(iv) A short section on education, health and productive health to capture potential impact on human capital investments and expenditures.

(v) A short credit section eliciting credit sources in the last year, the characteristics of credit contracts (with type of lender, loans requested and granted, duration, interest rates and

collateral requirements), credit history (with past applications for loans and rejected requests), and perceptions of the credit sectors by non-participants (incl. reasons for not applying and perceived likelihood to get a loan).

(vi) A section capturing shocks and risk coping and management behavior capturing the occurrence of different types of shocks (notably natural disasters, crop losses, deaths and sicknesses) over the last years, the subsequent losses and subsequent behavioral changes (incl. distress sales, dis-saving, increases in labor supply, reductions of human capital investments).

(vii) A section on attitudes towards the future, expectations and aspirations eliciting notably expected economic opportunities (e.g. perceived returns to given investments in agricultural assets or training) and aspired socio-economic mobility (e.g. aspirations with respect to the expansion of economic activities and to the education of children).

(viii) A section on environmental resources capturing the use and investments in specific local resources (incl. water, trees, energy sources), perceived benefits and costs of specific environmental investments and the presence and types of conflicts in the management of local environmental resources.

(ix) A section on the participation to and perception of the IDB/CIAT land tenure securisation program and of other government or NGO programs.

As mentioned above, three waves of data will be collected for the clarification stage, and the baseline survey will be fielded before any intervention, and notably the clarification of rights stage, takes place in the pilot areas. Information that allows identifying plots and linking the subsequent plot-level information from the plot surveys and program monitoring should thus be collected in this baseline survey. This will encompass the name of the locality (or "habitation") in which the parcel is situated, as well as a sketch(map) with the identification of the owners of the contiguous parcels (this is the method that Haitian private and public surveyors use for locating parcels). Once the clarification of rights data becomes available, the baseline survey dataset will be complemented with the identification code of the parcels defined by the program.

5.8 Data analysis strategy

As mentioned above, the analysis following the follow-up surveys for the clarification and formalization will allow estimating the short-run effects of the clarification and formalization of rights on outcomes of interest, by comparing the outcomes of individuals in the treatment and control as well as eventual indirect effects on the control group (using the comparison group). The basic econometric analysis at this stage will consist in estimating the following equation:

(eq1) $Yi = \alpha 1 + \beta 1\ T1i + \gamma 1\ Xi + \varepsilon 1i$

Where Yi are the outcomes of interest measured for household i, T1i is the treatment of the individual/household's parcels by the clarification of rights stage of the program, Xi a vector of controls for observable characteristics (including the baseline value of Yi).

The OLS estimate of β is a measure of the effects of the clarification of rights on the parcels.

The evidence on formalization will be obtained by comparing the outcomes of households that have been randomly selected as eligible for formalization with those of the control. The corresponding estimates will be obtained by estimating, for the sample of households incorporated in the formalization pilot:

(eq2) $Yi = \alpha 2 + \beta 2\ T2i + \gamma 2\ Xi + \varepsilon 2i$

Where T2i is being eligible for formalization.

The OLS estimate of $\beta 2$ is a measure of the effects of formalization (as intent to treat) on the outcomes of interest.

The endline analysis will provide estimates of the longer-run effects of the clarification of rights component by comparing the mean outcomes of the individuals in the treatment and non-experimental comparison group. The basic estimates that will be run at this stage are thus:

(eq3) $Yi = \alpha 3 + \beta 3\ T3i + \gamma 3\ Xi + \varepsilon i$

Where T3i is the treatment of the individual/household's parcels by the clarification of rights.

The OLS estimate of $\beta 3$ is a measure of the long-run effects of the clarification of rights component.

The stratification implicit in the pairwise matching procedure will allow estimating the heterogeneity of the impacts of the different phases of the program along different dimensions, in particular geographical area and proximity to urban centers.

6. Complementary Qualitative Evaluation

Each of the follow-up surveys will be preceded by a targeted qualitative evaluation. The objective of this qualitative evaluation will be:

(i) to document the implementation of the different stages of the intervention. This will serve to provide operational feedback to the implementing team to allow improvements in the implementation for the second group of localities. In addition, it will help to discover possible mechanisms underlying the quantitative findings.

(ii) to uncover qualitative perceptions and unanticipated impacts of the interventions. This will serve to prepare potential additions to the quantitative survey instruments. The piloting of these new additions will occur in concurrence with the qualitative survey.

The qualitative evaluation is envisioned to be a targeted evaluation, in which a small sample of treatment, control and comparison localities will be visited to interview producers and community leaders. In addition, key stakeholders in each comun will be interviewed (such as notaries, surveyors, judges, etc…). A short instrument with broad research questions will be prepared for this exercise, which will be conducted by a small team consisting of experienced enumerators and investigators. The list of questions will be updated through daily debriefings. The main product of these evaluations will be a short policy brief regarding implementation for direct operational purposes, a short summary of main findings, and the revised quantitative survey instruments.

7. Reporting Results

The results of this evaluation will be presented in a number of subsequent reports:

1. Analysis of baseline data to document initial characteristics of treatment, control and comparison groups, as well as the initial situation and perceptions regarding tenure status (type of documents (title), indivision, perceived security, and land conflicts.
2. Analysis of first follow-up survey to establish short-term impact of clarification of rigths (including possible spillovers) following methodology established above. This will include a short lessons learned from the qualitative evaluation.
3. Analysis of second follow-up survey to establish short-term impact of land formalization, as well as the longer-term impacts of clarification of rigths (including possible spillovers)

following methodology established above. This will include a short lessons learned from the qualitative evaluation.

These results will be shared and discussed on a continuous basis with the implementation team of CIAT and IDBs project team. In addition they will be disseminated to different stakeholders in Haiti, the IDB and academic audiences.

The proposed dates for the publication of the results will be determined based on the timing of the different project phases. Quantitative results will be published within a timeframe of 1 year after completion of each of the data collection rounds. Qualitative results will be shared with CIAT and the IDBs project team within a month of completing the qualitative evaluations.