Community Monitoring and Technology Transfer to Prevent Deforestation: A Field Experiment in the Peruvian Amazon
Pre-Analysis Plan

Jacob Kopas
Columbia University

Tara Slough
Columbia University

Johannes Urpelainen
Johns Hopkins SAIS

January 23, 2018

Abstract

Continued and accelerating deforestation of the Amazon rainforest represents a grave threat. While substantial investments have been made in satellite monitoring to detect deforestation events in the region, much of this data remains inaccessible to "first responders," namely indigenous communities in the Amazon and, to a lesser extent, the officials tasked with prosecuting these crimes. To what extent does providing this information to forest communities allow them to respond to such threats and better manage collective forest resources? Further, does facilitating contact between affected communities and government authorities promote increased legal action against offenders? We develop a Pre-Analysis Plan for a field experiment in the Peruvian Amazon that will provide indigenous communities with training, technology, and eventually weekly updated data detected by satellite monitoring on threats to collectively-held forests. A second treatment arm will share information corroborated by community monitor teams with state environmental prosecutors. We measure a rich set of outcomes using satellite data, original end line surveys, and administrative data on cases prosecuted to assess the impacts of the community monitoring program.

1 Introduction

This document provides a pre-analysis plan for the field experiment "Community Monitoring and Technology Transfer to Prevent Deforestation: A Field Experiment in the Peruvian Amazon." The goal of the pre-analysis plan is to prevent data mining and ensure that hypotheses are tested based on an analytical framework derived from first principles (Casey, Glennerster, and Miguel, 2012). This project represents one of the six Metaketa-III Natural Resources Governance studies. The results of this study will also be incorporated into a larger meta-analysis of the six projects.

Our primary intervention is an offer for an indigenous community to participate in a monitoring program of the community's collective forests. The monitoring program itself provides communities with monitoring technology (GPS smartphones), organizational support, and monthly-updated satellite imagery. Communities will then use the program to identify threats to collectively-held forests and take action to prevent or mitigate deforestation through community government or by informing national officials.

This study will also implement a secondary intervention in half of the treated communities. This intervention includes the filing of complaints about deforestation events to a national-level enforcement body in the department capital (Iquitos). Specifically, this treatment arm will solicit state intervention before the regional office of the national environmental prosecutor (*Fiscalía Nacional del Medio Ambiente*). In order to maximize power and understand the bureaucratic process underlying responses, we randomize the timing of these filings.

The main organization in charge of implementing the intervention with be Rainforest Foundation US in direct coordination with local indigenous federations and in consultation with the Peruvian Ministry of the Environment's National Program for Forest Conservation and Climate Change Mitigation (PNCBMCC) and other local partners. Indigenous federations are political bodies that unite indigenous communities in a geographical region and are run by elected officials from those communities. Federations have delegated authority to coordinate and propose projects or development plans with member communities, but which must be approved by a vote from community delegates. Rainforest US and the researchers will coordinate with the ORPIO federation, which has jurisdiction over the Loreto department, and the ORKIWAN and

FECONAMNCUA federations of the Napo basin, and the FECONATIYA and FECOTYBA federations of the Lower Amazon river basin.

The primary outcome indicator for this study will be the relative rate of deforestation in indigenous community territory. We will measure this outcome through Landsat satellite data provided by the University of Maryland Global Land Analysis and Discovery (GLAD) lab, and processed through the *Alertas Tempranas* (early warnings) system managed by Geobosques, a department within the Peruvian Ministry of the Environment. This study will also measure the impact of the intervention on community government and perceptions of community members through an end-line survey.

This intervention is a continuation of a preexisting program through Rainforest Foundation US, and which RFUS has already been successfully piloted in several communities of the Shipibo Conibo peoples in the Ucayali Department. Our partners conducted the pilot program in 2016 in a combined area of 20,000 hectares of titled indigenous lands that abut the Sierra del Divisor National Park, and which had suffered an increasing number of deforestation events according to the GLAD satellite alert system. RFUS, partnering with indigenous federations in the area, successfully implemented the project in the participating communities, and established monitoring teams which verified deforestation events and collected data. The Peruvian government eventually used this information as evidence to open an investigation into the deforestation on community lands. RFUS's successful implementation of this pilot project provides a firm basis of knowledge and experience that will permit our partner's to continue to implement the project in the selected river basins.

We are confident that the planned intervention will enjoy a relatively high level of compliance among communities in the Napo and Lower Amazon river basins, given the experience of RFUS with pilot communities and the shared features and similar circumstances between the pilot communities and the proposed intervention areas. Pilot communities and the subjects of this study are all threatened by deforestation, either currently possess a collective title for collectively-held lands or are in the process of obtaining one, and have large areas of forested communal lands, on average more than 10,000 HA. The indigenous federations that govern participating communities have also expressed interest in implementing the project within their jurisdiction.

2 Study Setting and Sample

2.1 Intervention Area

The selection of the intervention areas for this study is based on the criteria used by our implementing partner for prioritizing community monitor training. They selected several river basins within the Peruvian Amazon (roughly corresponding to the departments of Loreto, Ucayali, and Madre de Dios) that conformed to the following criteria:

1. Presence of recognized indigenous communities (i.e. *Comunidades Nativas* under Peruvian legislation) with the majority of their territory still forested and territory averaging 10,000 hectares.
2. Relative density of communities and ease of transportation
3. Lack of previous forest monitoring interventions
4. Principal environmental threat to communities is deforestation (i.e. no petroleum or large mining operations)
5. Community is not involved in the growing of coca.[1]
6. Political stability of indigenous federations (e.g. no foreseen elections or turnover in leadership) during project implementation period.

The river basins that fulfilled these requirements were the Lower Amazon (the portion east of Iquitos until the border with Brazil and Colombia) and the Napo basin (between the Ecuadorian border and Iquitos). Based on the above criteria, these river basins were selected as the two main intervention areas of the study.

[1]New varieties of coca have emerged that can be grown at substantially lower altitudes, including in the Amazon. Coca represents an illegal source of revenue for communities. We exclude coca growing communities as they are unlikely to accept participation in a monitoring program for obvious reasons.

2.2 Pre-selection process

We conducted a pre-selection process with federation leaders to determine the final sample of communities from which we will then make the randomized selection into the primary treatment arm. The pre-selection process provides two benefits for the study. First, the process will help identify the complete universe of indigenous communities in each of the selected river basins. This is necessary because although we are currently using the most accurate database of indigenous communities (which combines government and NGO collected data), this information is likely incomplete and there could be additional communities in each basin that could potentially be included in the study.[2] Secondly, a pre-selection process will allow us to gauge initial interest in the project and thereby limit the risk of non-compliance with the randomized selection process by communities that do not accept the offer to participate. Such non-compliance would require limiting our estimand of interest to the CACE, (Complier Average Causal Effect), that is, the average treatment effect among communities that are pre-disposed to accept treatment.

The pre-selection process took place in July and August 2017, and consisted of a trip by our implementing partners. During this trip, our partners identified a complete list of indigenous communities in each basin and meet with community leaders and federation leaders from each eligible community within four federations. The criteria we applied for determining the final subject pool is based on the fifth and sixth criteria for basin selection (lack of coca production and political stability), in addition to the federation leader's subjective perception of the likelihood of a community accepting the invitation to participate in the intervention.

After randomization, there will still be a chance for non-compliance, which we define as failure to accept and complete the monitoring training. Non-compliance will be identified after the first round of socialization meetings with communities selected into the treatment group. During these meetings, which will take place within each community or in a neighboring community (for communities in close proximity), our partners will describe the monitoring project and re-

[2]Data on indigenous communities come from IBC (2016) which combines data collected from regional governments, NGOs, and direct fieldwork. Interviews and consultations with our partners Rainforest US and other organizations and government officials in Peru have confirmed that this database represents the most accurate and complete collection of data on native communities in the Peruvian Amazon. This information is also used by national indigenous federations, NGOs, and the government in negotiation of land titling in the Amazon region.

quest that the leaders discuss the possibility within their communities and express their desire to participate through a letter.

2.3 Study Sample

We identified 73 eligible communities, of which 34 were assigned to treatment. We created 17 blocks, stratifying on the four federations (which implies stratifying on river basin), community titling status, and then area deforested in 2015 and 2016 (in bins by federation, titling status). From each block, we randomly selected two treatment communities and one alternate community. The random assignment was videotaped and conducted in the presence of the federation leaders in August 2017.

The community monitoring treatment is assigned at the community level. As per two of our outcomes, satellite measurements of 30×30 meter grid cells on a monthly basis and endline surveys, this study employs a cluster randomized design. Thus, all units of analysis are either community level or clustered by community. We will assign 20 of the communities in the Napo basin to the primary treatment of an offer to join the monitoring program. There are 23 control communities in this basin. We will assign 14 communities to this treatment and 16 communities to the control condition in the Amazon river basin. Thus, the total number of communities receiving the monitoring treatment is 34 and the number of control communities is 39.

The indigenous communities that will participate in this study consists of small villages of a few family clusters belonging to the same ethnic group, and having on average just under 300 members each. Communities in the intervention region will typically have one or two primary village settlements, although some have several, smaller satellite settlements of a few families within the territory. However, each community is organized around a single government structure consisting of an elected council (*Consejo Comunitario*) and a deliberative body (*Asemblea Comunitaria*), which is comprised of all members of the community and is the ultimate authority with state recognition and legal jurisdiction over the community's collective lands. Rainforest Foundation US, has contact with leaders from communities in both river basins, but has not yet begun to implement forest monitoring projects with communities in either area.

Although Peruvian legislation requires the national government to provide legal recognition and titles for all indigenous community lands, in practice titling had effectively halted after land reform efforts ended in the mid 1980s. However, the overwhelming majority of communities in each of the selected river basins have received titled land. In addition, a new wave of titling sponsored by the Inter-American Development Bank through the PTRT-3 (*Proyecto de catastro, titulación y registro de tierras rurales en el Perú*), will title 403 new indigenous communities, and will cover untitled communities within the project area. We consider the territory of the community for purposes of this study to equal the community's formally titled and/or demarcated land.

There exists a complication in demarcating community boundaries in untitled communities. While this would be completed during the training visits for treatment communities, we will not be able to measure the territory of untitled communities in control before the intervention. Because this affects only a small number of communities in the Amazon basin, there are three options for addressing this issue, in order of the steps that will be taken.

1. Meet with the DRA (*Dirección Regional de Agricultura*) of the regional government in Iquitos and obtain information info on the area demarcated for these communities

2. During endline data collection, ask community authorities in untitled control communities for their titling claims

3. Estimate the extent of the territory by using he following algorithm:

    - Calculate the average number of hectares for that Cuenca, and use that statistic as the estimated extension of territory;

    - Draw a polygon for which the community point is in the midpoint of one edge, and the rest of the territory extends away from the river to fill out the full etimated extension;

    - Other edges will stop at neighboring communities and rivers, or until they reach the full length of a side of a square with an area equal to the estimated area.

All communities that will participate in the study live in the Peruvian Amazon rainforest, one of the most biodiverse areas in the world, with the second highest number of bird species in the world and in the top 10 in nearly every other IUCN biodiversity category. The study area is a rich carbon sink, storing an average of 150 Mg C per hectare. Forest resources are also very important for sustaining indigenous communities' culture and economy. Peruvian Amazon indigenous communities are culturally and spiritually tied to their forests and traditionally depend on them for food security and medicine.

The communities participating in the study have collectively-held land that is largely forested. However, like many indigenous communities in the Amazon basin, these forest resources are threatened by unauthorized logging, non-indigenous colonizers using slash-and-burn agriculture, the cultivation of illicit drugs (coca), and illegal gold mining. These threats are highest during the dry season in the Peruvian Amazon, which corresponds roughly to the months of May through October, when timber is easier to cut and forests are easier to clear for agriculture. Although most communities are aware of and concerned about the general threat to their lands, effective protection of community forests are frequently hindered by three problems: (1) community lands cover a large area of forest, and specific information of individual events of deforestation are difficult to identify; (2) limited organizational or community resources often prevents communities from coordinating an effective response to deforestation threats; and (3) sharing information with national-level authorities with the power to enforce existing laws against unauthorized deforestation is difficult for remote communities. The Rainforest Foundation-US monitoring program specifically addresses these concerns.

The parties responsible for illegal clear-cutting and occupation of indigenous lands in the study intervention area are typically colonizers from Peru's costal or highland regions. These colonizers occupy areas of indigenous territory assuming that either the land is wild, state land (under Peruvian legislation, *tierras eriazas*, which can convert to private land after occupation and productive use), or that any occupation will be unnoticed by the indigenous community since indigenous territories tend to be large (frequently over 5,000 hectares) and community population sizes relatively small (100-200 families). In some instances, colonizers will enter into negotiations or rental agreements with select indigenous families for occupation of land or timber extraction. For these cases, the rental agreement is rarely if ever discussed with the community as a whole and can lead to tensions within the village, as the occupied land is considered collective property of the entire community. In both cases, the community often does not discover the occupation until colonizers are settled or logging and mining operations are well under way, at which point it becomes quite difficult for the community to eject the invading parties.

As mentioned above, the primary threat to forest resources is from clear-cutting due to agricultural colonization, illegal clear-cut logging, and gold mining. There is also less-invasive selective logging of highly valuable wood in community lands. We will not measure the impact of monitoring on this type of activity for two primary reasons. First, according to our partners, Rainforest US, most of the highly valuable timber in the intervention area has already been extracted, and this activity is not a major driver of deforestation. Second, the satellite data and the GLAD early alert system we rely on for both the monitoring treatment and for outcome measures are not sensitive enough to measure forest degradation from selective logging. Therefore we would not expect the monitoring treatment to have an impact on this type of activity, nor would we be able to measure the impact.

3 Treatment and Randomization

3.1 Primary Treatment Arm

The study has two treatment arms. This study's primary intervention is an offer for a community to participate in a monitoring program of the community's collective forests. The monitoring program itself provides communities with monitoring technology (GPS smartphones), organizational support, and training to use weekly updated satellite imagery to identify potential deforestation events. Communities will then by able to use information provided through monitoring to possibly take action to prevent or mitigate deforestation.

Our partners will implement the monitoring treatment over a period of three months from October to December 2017. The training will consist of three visits to each treatment community, although trainings in nearby communities will be combined to reduce time and travel costs. During the first visit, our partners will present information about forest monitoring program and develop a community monitoring plan. Each plan will identify key resources and threats according to the interests and existing knowledge of each community. The community will also identify a list of volunteers for participating in monitoring teams. Rainforest US will recommend that a monitoring team should consist of three monitors, and will provide enough equipment and payment for only three individuals.[3] These volunteers will then be trained in the use of GPS technology (smartphones) over the next two training dates. Teams will also be provided a monthly salary for the duration of the project (through December 2018), and also be given supplies and resources to patrol lands by boat and foot. During the final training period, the monitoring teams will conduct test patrols with trainers using recent satellite data. By February 2018, monitoring teams will be fully trained and begin able to conduct regular patrols.

The project will also provide a monetary compensation for monitoring teams of the equivalent of $80USD per monitor/month. According to consultations with our partner, Rainforest US, this payment is a necessary economic compensation for the individual labor monitors invest in the program, and the project would not be possible without this compensation. We will not include a separate monitoring arm without compensation.

Monitoring patrols will consist of two main activities. First, monitoring teams will use GIS data from satellite alerts to identify possible deforestation events on their territory. They will then visit these sites and identify and potentially document (taking GIS data, photos, etc.) the location and source of the threat. Information will be recorded using an intuitive and user-friendly app (Locus Pro Map) designed especially for GIS data collection. Second, monitoring teams will also be able to conduct regular patrols of territory and respond to potential threats identified by other sources (word of mouth, sightings by community members, etc.). Manuals for monitoring trainings developed by RFUS are on file with the EGAP project and available from researchers upon request.

In addition, our partners will conduct monthly visits to each treated community during the project implementation period. During these follow-up visits, our partners will provide updated satellite data from the Geobosques/GLAD early warning system through SD cards that can easily be downloaded into smartphones. This early-warning system is based on weekly updates

[3] However, the final number of monitors will be dependent upon the availability of volunteers and the autonomous decision of the community. Therefore, individual monitor teams may be composed of more or less members, who will share the equipment and payments accordingly.

of Landsat satellite imagery (30m resolution) analyzed through the University of Maryland and provided freely on the internet. Rainforest US will provide monitoring teams with a monthly summary of early warnings, which teams can then use to compare potential threats to community resources identified through the early-warning system with deforestation activity discovered first-hand through community monitoring. Communities will also provide our partners with update on the monitoring activities of communities and any responses to forest threats planned by the community.

Rainforest US will also use monthly trips to collect information on the frequency of patrols and any community meetings or decisions resulting from the monitoring intervention. Information from satellite data and monitoring teams will them be shared within the community. Colonizers could also potentially gain access to satellite information, but the data is also available freely on the internet to any colonizer with an internet connection.

3.2 Secondary Treatment Arm

The secondary treatment arm will test the effect of providing information to and engaging with environmental prosecutors (*Fiscalía Especializada en Materia Ambiental*, FEMA), national-level authorities that can mobilize state resources for the protection of community forests. This intervention is a rollout of the filing of official complaints (*denuncias*) by community members. Environmental prosecutors are part of the national government and charged with enforcing environmental protection laws and persecuting environmental crimes, which include illegal logging and deforestation on indigenous community lands.

Specifically, project partners will work with community members to submit an official complaint with local FEMA offices in Iquitos. The complaint will communicate information collected through community-level monitoring and through the early-warning satellite data regarding specific threats identified in community forests. After receiving information from community complaints, environmental prosecutors will then decide whether to take further action regarding each community. FEMA environmental prosecutors are authorized under Peruvian law to open investigations concerning illegal deforestation, impose sanctions, and start criminal prosecutions.

This treatment arm will be randomized across communities that have been assigned to the

River	Federation	Block 1	Block 2	Block 3	Block 4	Block 5	Block 6
Napo	ORKIWAN	T, $n = 5$	T, $n = 4$	T, $n = 4$	T, $n = 4$	T, $n = 4$	T, $n = 4$
Napo	FECONAMNCUA	NT, $n = 5$	T, $n = 4$	T, $n = 5$	T, $n = 4$		
Amazon	FECONATIYA	NT, $n = 4$	M, $n = 4$	T, $n = 5$			
Amazon	FECOTYBA	NT, $n = 3$	NT, $n = 4$	T, $n = 5$	T, $n = 5$		

Table 1: Blocks. Among titled and untitled blocks, blocks were created based on past deforestation levels. The abbreviations T, NT, and M indicate titling status, titled, not titled, and mixed, respectively.

River Basin	Control Communities	Monitoring only	Monitoring + State Engagement
Napo	43	10	10
Amazon	30	7	7

Table 2: Allocation of communities to each treatment arm.

primary treatment condition only. The randomization will be blocked so that an equal number of communities are selected into the secondary treatment arm in each basin. Thus, the primary intervention in a three treatment arm design that includes: (1) control communities; (2) monitoring-only communities; and (3) monitoring and state engagement communities. Within the units assigned to the third treatment arm, the complaints will be rolled out over time through a second randomization process. This allows us to increase power in this arm and test the process of enforcement in the FEMA office.

3.3 Random Assignment

In collaboration with indigenous federation leaders, we created seventeen blocks of three to five units. We blocked completely on federation (and thus river basin), titling status (with one mixed block), and then on deforestation in the past two years. The composition and size of the blocks are listed in Table 1. As is evident from the table, we will block by river basin. We also will seek additional efficiency gains by blocking within each river basin on geographic characteristics and past deforestation patterns (2012-2017).

Table 2 summarizes the treatment assignment. Within each block, we will assign one unit to the "monitoring only" treatment and one unit to the "monitoring + state engagement" treatment. Given the heterogeneity in the size of these blocks, all empirical specifications will use IPW.

The randomization was conducted in August 2017 upon conclusion of the community selection process. Prior to community selection, RFF will create a list of pre-selected communities together with indigenous federation leaders. We assume that this will minimize the threat of non-compliance. If, however, a treatment community refuses to participate/complete training activities, there is a randomly assigned "backup" community from each basin that will be assign treatment in the case of non-compliance by one of the assigned communities. Both communities (original and replacement) will be coded as "assigned to treatment" for the purposes of analysis if they are assigned to treatment.

4 Implementation Plan

The study will be implemented as follows:

1. **July-August 2017**. Pre-selection of communities; approval by community-level and regional indigenous authorities; recruitment of local staff; primary treatment arm assignment.

2. **September-December 2017**. Implementation of community monitoring trainings for the primary treatment arm. Community monitoring teams will be equipped and fully trained by the end of this period.

3. **January-February 2018**. Community monitoring teams patrol community forests in treated communities. A second treatment arm—information sharing with national authorities—will be rolled out during the period on a staggered basis. The research team will measure deforestation with satellite data.

4. **November-December 2018**. End line survey in treatment and control communities.

5 Theory and Hypotheses

5.1 Theory: Information, Coordination, and State Engagement

The community forests under examination in this study are examples of Common Pool Resources, or resources that are sufficiently large to make it impossible to fully exclude others from use, but which can theoretically be managed so that use of the resource does not undermine the longterm stock (Ostrom, 1990). The extensive literature on Common Pool Resources governance has highlighted the importance of resource monitoring in contributing to the emergence and maintenance of local institutions that can solve collective actions problems and promote the sustainable use of the resource (Ostrom, 1990; Dietz, Ostrom, and Stern, 2003). The primary mechanism through which monitoring can contribute to sustainable resource governance is through the provision of information to the users of the resource. By decreasing costs on information, users can rationally adjust their behavior and coordinate to sanction non-compliance by other users.

However, information may not be the only mechanism through which the intervention can affect forest management. The intervention itself provides training and leaves an organization structure that can also help promote collective action within the community, and in particular around problems related to management of community forests. This increased organizational capacity, which includes increasing human capital within the community, can also increase the community's ability to mobilize politically and demand better enforcement by the state or to enforce community decisions against external actors. Increasing organizational capacity and human capital have also been signaled as key to the development of the indigenous people's movement more broadly in Latin America (Yashar, 2005).

Finally, the exact effectiveness of local institutions in controlling and managing common pool resources is likely dependent on variation in local conditions, including the specific characteristics of the resource and users in question (Agrawal, 2007). For the communities within our sample, the primary users (community members) have limited involvement with direct drivers of deforestation. Rather, the primary threats to forests resources are external actors that are not involved with community institutions that regulate resource use. Community government can police cooperation of members by limiting unauthorized contracts or deals with external actors, and by mobilizing community participation to enforce sanctions against non-community members. However, given this context, external enforcement by state actors may be necessary for controlling deforestation and excluding external actors from community lands.

Our study will examine each of these three above mechanisms: information, coordination, and engagement with the state. The first mechanism through which the present treatment works is by providing communities with recent information about deforestation. While satellites and Geobosques' *Alertas Tempranas* system has been identifying deforestation weekly since since at least 2012, this information is not readily accessible to indigenous communities in remote parts of the Amazon forest. Therefore, the intervention's use of smartphone technology and monthly visits can provide treatment communities with remote-sensing satellite and direct observational data that would not otherwise be available. Since affected communities serve as the first line of defense against illegal deforestation for wood or agriculture, information transmission should reduce "search costs" (i.e. costs related to obtaining and processing information) for relevant information on resource use and status.

Second, the training and monitoring program aims to strengthen communities' response to threats to their communal territory. The information provided via the monitoring system and the work of the community monitors should allow communities to coordinate around specific threats to forest resources. From an organizational perspective, the hiring of a monitor also promotes a more centralized process in terms of responding to deforestation activities. Coordinated responses to threats to communal lands should, in principle, provide a stronger line of defense against illegal logging. As such, changes in coordination capacity and community organization with specific reference to deforestation should also reduce the extent of deforestation activity within their territories.

Finally, the Peruvian state ultimately has responsibility for enforcing its laws with regard to deforestation. Such enforcement is quite costly in the river basins that we study. First, these are peripheral areas where the state has typically had a limited presence. As such enforcement is quite costly, particularly in the absence of information about deforestation activity in the form of complaints (*denuncias*). Second, the indigenous communities in which the intervention will be conducted typically have limited contact with the state. There are many explanations for this limited contact ranging from racism, cultural/language barriers, and limited state presence. By empowering communities to bring their grievances to state agents, the probability of state enforcement against offenders should increase. The threat of such enforcement should deter continued deforestation.

Upon receiving information on the threat of deforestation on community lands, the community can take one of two (or both) courses of action: (1) inform state officials; or (2) take autonomous action as a community. In either course of action, the community will act through its two principal governing bodies, the village council (*Consejo Comunitario*) and the village assembly (*Asamblea Comunitaria*). The village council is an executive board of elected officials (all community members) that initiate actions in name of the community. However, the village assembly, which consists of all members of the community, is the highest authority in the community and must approve all community decisions. The village council will propose a given course of action, and the assembly will then vote to approve.

We will measure the impact of the first course of action—informing state officials—through our secondary treatment arm mentioned above (Section 3.2). The secondary course of action can involve many different decisions, all of which are within the power and authority of the village councils and assemblies to make regarding their territory. While it will be impossible to predict the exact actions taken in each case, typical actions in the past involve placing control points in key entry paths (e.g. along rivers or streams connecting the community to urban centers), fining members who negotiate with colonizers without community approval, or directly confronting colonizers. Given the lack of state presence in community lands, it is possible that any of these course of action entail increased risk to the safety of community members. Rainforest US includes in its training protocol methods to reduce the chance of risk, which we outline below in Section 12. However, it is important to note that these decisions are completely within the legal authority of the community, and the community ultimately has sovereignty to govern its territory, and therefore neither the researchers nor Rainforest US can dictate the actions the community should or should not take.

Via these three mechanisms, we posit a set of hypotheses linking the community monitoring intervention to lower levels and rates of deforestation in treatment communities. We specify a larger set of outcomes in order to test the mechanisms underlying the predicted relationship with deforestation.

5.2 Main Hypotheses, Satellite Data

Hypothesis 1. ***Main Effect of Monitoring:*** *Relative to the control group, community monitoring (both treatment arms) reduces deforestation events and area.*

Hypothesis 2. ***Marginal Effect of State Engagement:*** *Relative to community monitoring only, the combination of community monitoring and information sharing with national authorities reduces deforestation events and area.*

5.3 Hypothesized Mechanisms and Heterogeneity, Satellite Data

Hypothesis 3. ***Logging Roads:*** *Relative to the control group, community monitoring (both treatment arms) reduces the construction of new logging roads in the communities.*

Hypothesis 4. ***Seasonality:*** *The negative effect of community monitoring (both arms) during the rainy season (approximately May-October) is stronger than during the dry season (approximately November-April).*

Hypothesis 5. ***Titling:*** *The negative effect of community monitoring is stronger in titled communities (at the time of randomization) than in untitled communities. [Our ability to test this hypothesis is conditional on having a sufficient number non-titled communities in the sample.]*

Hypothesis 6. ***Population:*** *The negative effect of community monitoring is stronger in communities with smaller population sizes, as they will better be able to resolve collective action problems.*

5.4 Hypothesized Spillovers, Satellite Data

Hypothesis 7. ***Displacement of Deforestation:*** *Relative to the control group, community monitoring (both arms) increases deforestation events and area within a buffer zone outside the monitored communities.*

Hypothesis 8. ***Displacement of Deforestation 2:*** *Relative to community monitoring only, the combination of community monitoring and information sharing with national authorities increases deforestation events and area within a buffer zone outside of monitored communities.*

Hypothesis 9. ***Displacement of Logging Roads:*** *Relative to the control group, community monitoring* *(both arms) reduces the construction of new logging roads within a fixed buffer zone.*

5.5 Main Hypotheses, Administrative Data Outcomes

Hypothesis 10. ***Environmental Prosecutors:*** *Relative to the control group, prosecutors should open a higher number of new cases in communities with community monitoring (both treatment arms).*

Hypothesis 11. ***Environmental Prosecutors and State Engagement:*** *Relative to the control group, prosecutors should open a higher number of new cases in communities with state engagement and community monitoring than with community monitoring alone.*

5.6 Survey and Community Governance Outcomes: Mechanisms

Hypothesis 12. *Awareness of Deforestation: Relative to members of control group communities, members of monitoring (both arms) communities manifest greater awareness of the incidence of deforestation in the Peruvian Amazon.*

Hypothesis 13. *Trust in the State: Relative to members of control group communities, members of monitoring (both arms) communities report higher or lower trust in the Peruvian state to address deforestation. [Two-sided]*

Hypothesis 14. *Community Governance: Relative to members of control group communities, members of monitoring (both arms) communities report more frequent and/or recent community meetings.*

Hypothesis 15. *Coordination: Relative to members of control group communities, members of monitoring (both arms) communities manifest higher levels of coordination in a set of lab games.*

6 Data Collection and Operationalization

Our outcome data comes from four distinct sources. We outline the characteristics of each data source in this section.

6.1 Satellite Data

The primary outcome measure for this project will be deforestation rates and area deforested on community lands. To measure deforestation, we will use monthly measures of Landsat satellite data provided by the University of Maryland Global Land Analysis and Discovery (GLAD) lab. Satellite imagery for the entire Peruvian Amazon is available publicly through the GLAD lab's website and is taken at a resolution of $30\times$ 30m grid cells. At this level of resolution, we will be able to detect the clearing of a forested grid-cell, although smaller levels of degradation (for example the selective logging of valuable timber such as mahogany) will not be detectable. The advantage of using this data to measure our primary outcome is that it is available freely and on a weekly-updated basis, there are satellite images available from several years prior to the start of the project, and data can be drawn for several years following the project, allowing us to measure the longer term impact of the monitoring project. The data typically identifies deforestation *events*, providing coordinates for each event. As such, our outcome measures are coded such that higher values correspond to more deforestation.

Once all communities are mapped in GIS, we will overlay the grid of $30 \times 30m$ cells from Landsat. All cells containing any community territory will be included in the experimental sample. (Avelino, Baylis, and Honey-Rosés, 2016) provides some commentary on the spatial aggregation of satellite imagery data. We plan to analyze at the most disaggregated level (30×30m grid cells). In contrast to much existing work, we know the level of treatment assignment, the community, so all standard errors will be clustered at this level.[4] The authors demonstrate that "the estimate of the treatment effect is unbiased when the true level is perfectly divisible by the unit of analysis" (15). The communities contain a very large ($\geq 10^4$) number of units so any bias resulting from the disaggregate unit of analysis is minimal.[5]

The satellite data also allows us to exploit temporal variation in deforestation activity. While there is in principle weekly satellite data, when there is cloud cover during image taking, however, there is additional noise in the data. As such, we aggregate at the monthly level to minimize the probability of obtaining no clear image within a time period, t. The temporal aggregation is uncorrelated with treatment assignment and provides a comprehensible unit of analysis. The "robustness" of results to this temporal unit can be assessed by moving the temporal window

[4]The efficiency results in Avelino, Baylis, and Honey-Rosés (2016) are driven mostly by the lack of information about the process of assignment to treatment. Such concerns are less relevant when this process is known.

[5]The Monte Carlo results in Avelino, Baylis, and Honey-Rosés (2016)suggest that under the specified DGPs, any bias would attenuate effects.

forward or backward as necessary. If there is no clear image of a given subset of grid cells in a given month, we impute a "0." Given the empirical strategy defined here, our main estimand is the ATE on monthly rates of deforestation. The estimator will average across months. Given that deforestation (in the short run) is a monotonic process, deforestation in months with cloudy images will be picked up on subsequent images, and the average across months is still estimable. This does add some noise to efforts to detect seasonal trends.

Finally, we will also use satellite data to measure the increase in new, informal road networks in indigenous collective lands. Together with rivers, these informal roads, which are often measurable with remote sensing data, are the principal means of extracting timber from community lands. This measurement will be contingent upon identifying a suitable methodology for identifying and measuring road network lengths, which we will research at a future date.

6.2 Survey Measures

The survey consists of an endline survey of 10 respondents per community. We will randomly sample one two control communities from each sub-river basin block.[6]. The survey intends to measure the theorized mechanisms underlying observed treatment effects. The survey will consist of questions and several behavioral games. These measures include:

- Awareness of the prevalence of deforestation and its effects.
- Awareness of community rights with respect to community resources.
- Awareness of legal provisions against deforestation.
- Trust in the Peruvian government
- Coordination games [tbd]

We also seek to conduct a much longer interview with the monitor in each community during this survey. The challenge, here, is identifying the individual that *would* have been the monitor in control communities. One option is to register a monitor candidate when gauging initial interest in the program during community selection.

[6] In the block with 3 units, we will include that control community in the survey

6.3 FEMA Cases

Our final outcome measures the new cases opened for environmental crimes. In principle, with more information on deforestation activity and resources to travel to state capitals to meet with the prosecutors, a larger proportion of cases should be reported to state authorities, and therefore a larger number of new cases or investigations opened by the national environmental prosecutor, the *Fiscalía Especializada en Materia Ambiente* (FEMA). Ponce et al. (2015) documents the availability of such case files in the offices of the FEMA. We propose to measure the number of new investigations opened regarding an indigenous community over the length of the study period. There are some limitations to this measure. First, since filing a complaint is conditional upon another post-treatment outcome, the potential outcome for the number of new investigations opened is undefined for deforestation events that did not happen. There are some gains to be made by aggregating events at the community level, but the interpretation of the estimand is somewhat different.[7]

Note that opening of an investigation does not imply that the state is effective in resolving the claim or bringing a case before a national court. Ponce et al. (2015) implies claim resolution rates of 24% in Ucayali and 41% in Loreto over the four year period from 2010 to 2014. Studying case resolution adds an additional layer of selection, so we do not propose to study this as an outcome.

6.4 Community Governance

We will also collect data on the change in government capacity of indigenous village councils and assemblies, the primary government bodies for indigenous communities in the Peruvian Amazon. Specifically, will gather data on community records of indigenous government in Peru; a source of information that has not yet been studied, but which has the potential to provide high-quality, time-series data on community governance.

Nearly all recognized indigenous communities in the Peruvian Amazon region follow a similar government structure established by national legislation and promoted by indigenous fed-

[7] We propose to estimate the bounds on the equivalent of a Survivor Average Causal Effect (SACE) as an alternative strategy to resolve some of the identification challenges present in the proposed measure.

erations to improve governance and access to national and international development aid. This structure consists of an elected executive body, the *Consejo Comunitario* (village council), and a deliberative body, the *Asemblea Comunitaria* (village assembly), comprised of all adult members of the community. Communities will typically keep a record of all important meetings of the council and the assembly through a series of short meeting minutes commonly referred to as *Actas Communitarias* (community registry). These registries of government activity will be the primary measure of organizational capacity. Increases in the number and frequency of community registries following monitoring should be an indication that the intervention also increases a community?s organizational capacity. Since these are historical records, the community registry will provide pre-treatment, baseline data with which to measure the change in community governance over time.

6.5 Community-level Covariates and Mediating Outcomes

We will also collect data on population size to test for conditional treatment effects for community size. This information will be collected during the initial socialization meetings with treatment communities and during the endline survey with control communities. Although this will involve using some post-treatment data, population data, which tends to be relatively stable in sample communities for short time periods such as the length of this study (1 year), is highly unlikely to be affected by the treatment. In addition, the study will collect data on mediating outcomes to provide evidence on SUTVA assumptions and that the treatment is producing an effect through the theorized mechanisms. Specifically we will collect information about the number of monitoring trips conducted during the study period and the number of deforestation events verified by monitoring teams in each community. We will define other community-level covariates to use for verifying randomization assumptions and improve accuracy of effect estimates at a later date. These may include population size, area of collective lands, ethnicity, and distance from Iquitos.

7 Estimation

We propose the use of several estimators given the characteristics of the data sources and the different hypotheses being tested. We divide these estimation strategies by outcome variable as follows.

7.1 Satellite Data

7.1.1 Effect on Communities in Sample

Let i denote grid cells clustered under communities j. Let t denote a month in year k. We denote the monitoring treatment as Z_j^M and the state engagement treatment as Z_{jt}^R. Z_{jt}^R is indexed by time and takes the value "1" in all months subsequent to the submission of the first denuncia. Note that communities have different amounts of territory, thus cluster size is heterogeneous. The number of clusters here, ≥ 100 should mitigate the possibility for bias on this account. We weight observations by the inverse of the share of observations We propose two estimators, each of which accounts for the long time series of pretreatment data that we have.

For the post-treatment period outcomes only, we estimate the following equation via OLS with IPW to account for differences in probability of assignment to treatment.

$$Y_{ijtk} = \beta_0 + \beta_1 Z_j^M + \beta_2 Z_j^{Rt} + \gamma \sum_{k=2012}^{2017} Y_{ijt} + \epsilon_{ijt} \qquad \text{for } k = 2018 \tag{1}$$

Here we adjust for pre-treatment grid-cell month outcomes from 2012 to 2017. Standard errors are clustered at the community level (j). This can easily be adapted to estimate treatment-by-covariate interactions. The prespecified covariates upon which heterogeneous treatment effects are estimated are measured at the community j or month t level. The X_j in Equation 2 could also be written X_t:

$$Y_{ijtk} = \beta_0 + \phi_1 Z_j^M + \phi_2 Z_{jt}^R + \zeta_1 Z_j^M X_j + \zeta_2 Z_j^R X_j + \delta X_j + \gamma \sum_{k=2012}^{2017} Y_{ijt} + \epsilon_{ijt} \qquad \text{for } k = 2018 \tag{2}$$

Alternatively, we will estimate a generalized difference-in-difference specification using the full panel as follows:

$$Y_{ijt} = \alpha_i + \kappa_{ty} + \beta_1 Z_j^M \mathbb{I}[k = 2018] + \beta_2 Z_{jt}^R \mathbb{I}[k = 2018] + \epsilon_{ijt} \tag{3}$$

Here, the indicator function $\mathbb{I}[k = 2018]$ indicates the post-treatment period. Again, clustered standard errors will be clustered at the community level (j). To estimate treatment-by-covariate interactions on an arbitrary pre-treatment X_j (or X_t), we estimate:

$$Y_{ijt} = \alpha_i + \kappa_{ty} + \mathbb{I}[k = 2018](\phi_1 Z_j^M + \phi_2 Z_j^R + \zeta_1 Z_j^M X_j + \zeta_2 Z_j^R X_j) + \epsilon_{ijt} \tag{4}$$

These specifications provide one basic and principled way to estimate temporal patterns in the outcome data and to exploit the randomized rollout of the second treatment arm.

7.1.2 Effect Outside of Communities in Sample

We also intend to estimate the possibility of spillovers outside treated and untreated communities. In this analysis, we again denote denote grid cells as i clustered by experimental community j. Let t denote a month in year k.

The sample for these estimates is a buffer zone within a fixed distance of the boundary with a treated community. Thus, the unit at which the cell is clustered is with the experimental community in whose buffer the unit is located. Analysis of pre-treatment data will allow for an appropriate specification of an appropriate buffer region and ensure that buffer regions are not overlapping.

We denote the grid cells with monitoring treatment as Z_j^M and the state engagement treatment as Z_{jt}^R. All estimation takes the same form as Equations 1-4. However, an additional distance term (linear distance from center of grid cell to the boundary of the experimental community) will be interacted with treatment to assess the spatial dynamics of spillovers.

7.2 Survey Outcomes

For the survey outcomes, we have multiple individuals, indexed i, per community j. The ATEs can be be estimated straightforwardly by estimating the following specification via OLS with

IPW to account for differences in probability of assignment to treatment within blocks.

$$Y_{ij} = \beta_0 + \beta_1 Z_j^M + \beta_2 Z_j^R + \epsilon_{ij} \tag{5}$$

Individual-level or community-level treatment-by-covariate interactions can be estimated by Equation 6. Note that the X_i represents an individual-level covariate but could equivalently be written for a community-level covariate.

$$Y_{ij} = \beta_0 + \phi_1 Z_j^M + \phi_2 Z_j^R + \zeta Z_j^M X_i + \zeta Z_j^R X_i + \delta X_i + \epsilon_{ij} \tag{6}$$

Clustered errors will be clustered at the community level (j), the level of treatment assignment.

7.3 FEMA Cases

Here, the outcome, C_j is the number of new cases opened by the FEMA prosecutors regarding the territory of community j during the year of 2018. There are two ways to measure this outcome, the raw number of cases opened C_j or the share of cases opened C_j relative to the total number of deforestation events E_j. This can be estimated quite simply using the following OLS specification with IPW:

$$C_j = \beta_0 + \beta_1 Z_j^M + \beta_2 Z_j^R + \gamma \mathbf{X}_j + \epsilon_j \tag{7}$$

$$\frac{C_j}{E_j} = \beta_0 + \beta_1 Z_j^M + \beta_2 Z_j^R + \gamma \mathbf{X}_j + \epsilon_j \tag{8}$$

Here, heteroskedasticity-robust standard errors will be estimated. $\mathbf{X}_j$ represents a matrix of community-level pre-treatment covariates (to be defined).

We will also bound an event-level Survivor Average Causal Effect (SACE) to address concerns about the endogeneity of the denominator in 8. Here "survivor" refers to events that would have occurred with or without a given treatment. In principle, unless treatment effects on deforestation are negligible, the assumption-free bounds on the SACE will be quite large. We will make a monotonicity assumption to narrow the bounds and examine sensitivity to relaxation of this assumption.

8 Mapping Hypotheses to Empirical Tests

We map from our hypotheses onto the estimators described above in Table **??**. The full set of covariates utilized in covariate adjustment remains to be defined. Furthermore, some outcomes such as the logging roads indicator are not yet well operationalized. Analysis of pretreatment data and patterns will inform the further development of this table in the coming months.

8.1 Main Outcomes

The estimands and estimators that we will employ to test the main hypotheses are listed in Table 3. We use two sets of covariates across different outcomes:

- **General [G]**: Block fixed effects
- **Lagged Dependent Variable [L]**: Lagged dependent variable, averaged (by month) from 2012-2017
- **Demographic Controls [D]**: Respondent gender, age, and age^2, and interviewer fixed effects

Hypothesis	Estimand	Equation	Sample	DV	Estimator	SE	Covariates	Test Type
1	ATE (ITT)	1	Exp. Communities	# of Events	β_1	Clustered	G, L	Lower
1	ATE (ITT)	1	Exp. Communities	Area Deforested	β_1	Clustered	G, L	Lower
1	ATE (ITT)	3	Exp. Communities	# of Events	β_1	Clustered	G, L	Lower
1	ATE (ITT)	3	Exp. Communities	Area Deforested	β_1	Clustered	G, L	Lower
2	ATE (ITT)	1	Exp. Communities	# of Events	β_2	Clustered	G, L	Lower
2	ATE (ITT)	1	Exp. Communities	Area Deforested	β_2	Clustered	G, L	Lower
2	ATE (ITT)	3	Exp. Communities	# of Events	β_2	Clustered	G, L	Lower
2	ATE (ITT)	3	Exp. Communities	Area Deforested	β_2	Clustered	G, L	Lower

Table 3: Prespecified empirical tests to test our main hypotheses. Covariate codes are listed above.

8.2 Mechanisms and Heterogeneity with Satellite Data

The estimands and estimators that we will employ to test the hypotheses on mechanisms and heterogeneity are listed in Table 4.

Hypothesis	Estimand	Equation	Sample	DV	Estimator	SE	Covariates	Test Type
3	ATE (ITT)	1	Exp. Communities	TBD (roads)	β_1	Clustered	G, L	Lower
3	ATE (ITT)	3	Exp. Communities	TBD (roads)	β_1	Clustered	G, L	Lower
4	CATE	2	Exp. Communities	# of Events	ϕ_1, ζ_1	Clustered	G, L	Lower (ζ_1)
4	CATE	2	Exp. Communities	Area Deforested	ϕ_1, ζ_1	Clustered	G, L	Lower (ζ_1)
4	CATE	4	Exp. Communities	# of Events	ϕ_1, ζ_1	Clustered	G, L	Lower (ζ_1)
4	CATE	4	Exp. Communities	Area Deforested	ϕ_1, ζ_1	Clustered	G, L	Lower (ζ_1)
5	CATE	2	Exp. Communities	# of Events	ϕ_1, ζ_1	Clustered	G, L	Lower (ζ_1)
5	CATE	2	Exp. Communities	Area Deforested	ϕ_1, ζ_1	Clustered	G, L	Lower (ζ_1)
5	CATE	4	Exp. Communities	# of Events	ϕ_1, ζ_1	Clustered	G, L	Lower (ζ_1)
5	CATE	4	Exp. Communities	Area Deforested	ϕ_1, ζ_1	Clustered	G, L	Lower (ζ_1)

Table 4: Prespecified empirical tests to test our mechanisms and heterogeneous treatment effect hypotheses. Covariate codes are listed above.

8.3 Spillovers

Given that the Peruvian Amazon region has vast areas of land, only some of which is occupied or demarcated by indigenous communities, it is likely that increased monitoring will not eliminate deforestation in the broader region. Rather, we expect that the primary effect of the monitoring treatment will be to displace colonization and deforestation to non-intervention areas or other regions. Finally, we do not expect the treatment to cause colonizers to shift clear-cutting activity to selective logging, as the two activities follow different economic patterns[8] and there is insufficient supply of highly valued timber for selective logging to be economically viable for colonizers.

The estimands and estimators that we will employ to test the hypotheses on spillovers using the satellite data are listed in Table 5.

Hypothesis	Estimand	Equation	Sample	DV	Estimator	SE	Covariates	Test Type
6	ATE (ITT)	1	Buffer Zone	# of Events	β_1	Clustered	G, L	Upper
6	ATE (ITT)	1	Buffer Zone	Area Deforested	β_1	Clustered	G, L	Upper
6	ATE (ITT)	3	Buffer Zone	# of Events	β_1	Clustered	G, L	Upper
6	ATE (ITT)	3	Buffer Zone	Area Deforested	β_1	Clustered	G, L	Upper
7	ATE (ITT)	1	Buffer Zone	# of Events	β_2	Clustered	G, L	Upper
7	ATE (ITT)	1	Buffer Zone	Area Deforested	β_2	Clustered	G, L	Upper
7	ATE (ITT)	3	Buffer Zone	# of Events	β_2	Clustered	G, L	Upper
7	ATE (ITT)	3	Buffer Zone	Area Deforested	β_2	Clustered	G, L	Upper
8	ATE (ITT)	1	Buffer Zone	TBD (roads)	β_1	Clustered	G, L	Upper
8	ATE (ITT)	1	Buffer Zone	TBD (roads)	β_1	Clustered	G, L	Upper

Table 5: Prespecified empirical tests to test our spillovers hypotheses. Covariate codes are listed above.

8.4 Administrative Data

The estimands and estimators that we will employ to test the hypotheses on the administrative outcomes are listed in Table 6.

Hypothesis	Estimand	Equation	Sample	DV	Estimator	SE	Covariates	Test Type
10	ATE (ITT)	7	Exp. Communities	Cases Opened	β_1	Huber-White	G	Upper
10	ATE (ITT	8	Exp. Communities	Cases Opened/Events	β_1	Huber-White	G	Upper
11	ATE (ITT)	7	Exp. Communities	Cases Opened	β_2	Huber-White	G	Upper
11	ATE (ITT	8	Exp. Communities	Cases Opened/Events	β_2	Huber-White	G	Upper

Table 6: Prespecified empirical tests to test our hypotheses from the administrative data. Covariate codes are listed above.

[8]The former resulting from migration from the highlands, while the later are smaller business operations from well-connected locals.

8.5 Survey Outcomes: Mechanisms

The estimands and estimators that we will employ to test the hypotheses on mechanisms from endline survey data are listed in Table 7.

Hypothesis	Estimand	Equation	Sample	DV	Estimator	SE	Covariates	Test Type
12	ATE (ITT)	5	Respondents in EC	TBD (survey)	β_1	Clustered	G, D	Upper
13	ATE (ITT)	5	Respondents in EC	TBD (survey)	β_1	Clustered	G, D	Upper
14	ATE (ITT)	5	Respondents in EC	TBD (survey)	β_1	Clustered	G, D	Two-Tailed
15	ATE (ITT)	5	Respondents in EC	TBD (survey)	β_1	Clustered	G, D	Upper

Table 7: Prespecified empirical tests to test our hypotheses from the administrative data. Note that the covariates to be included in each specification are yet to be determined.

8.6 Multiple Comparisons

Clearly we intend to test multiple relationships and hypotheses as manifest in our hypotheses and Tables 3 to 7. As such, we aim to address the threat of multiple comparisons in one of two ways. The first approach follows Kling, Liebman, and Katz (2007) by dividing our outcomes into families and assessing the effects on a standardized index measure from each family. The second approach follows Benjamini and Hochberg (1995) and implements a false discovery rate (FDR) correction.

9 Power Analysis

Given the multiple sources of outcome data with differing units of analysis, we present the results of power simulations on two conservative estimators we employ. These simulations assume full compliance with treatment assignment and no attrition. Earlier in the project, we characterized and analyzed our design in `DeclareDesign` and will share these analyses upon request. The analysis below is a reduced version of those simulation.

9.1 Survey Outcomes

In the context of survey outcome data, we assume 10 respondents per community. We vary the ICC as well as the ITT (here ATE) size. We estimate a regression without covariate adjustment following Equation 5, focusing on the estimator β_1, the effect of monitoring. We report the power of both one- and two-tailed tests at the $\alpha = .05$ level in Table 8.

	Simulated Standardized ATE				
ICC	0.1	0.2	0.3	0.4	0.5
0.00	0.236; 0.324	0.624; 0.728	0.912; 0.94	0.99; 0.996	1; 1
0.20	0.112; 0.202	0.27; 0.396	0.522; 0.622	0.766; 0.854	0.916; 0.958
0.40	0.082; 0.122	0.206; 0.276	0.382; 0.47	0.568; 0.688	0.724; 0.82
0.60	0.098; 0.136	0.154; 0.242	0.266; 0.368	0.396; 0.53	0.556; 0.704

Table 8: Power for one-tailed tests (left and right, respectively) using simulated survey data without covariate adjustment. Each cell represents the estimate over 500 simulated data sets.

Table 8 suggests that for moderate ICC values ($\approx .2$), we can detect a sizable standardized effect of 0.4. Covariate adjustment should increase the precision of the estimates and may lead to modest power gains.

9.2 Deforestation Data

We do a power calculation that mimics our approach to the satellite data with several simplifying features. First, we assume the community (not the grid cell) is the unit of analysis, which should be conservative. In this context, we assume the dependent variable to be continuous representing, for example, the share of area deforested in a given month. We further assume a seasonal pattern, derived from existing deforestation data from the Amazon but standardize by month, such that the simulated treatment effect is a standardized effect on the share of area deforested in a given month. Variance in potential outcomes can be decomposed into the unit, community, month, and observation (e.g. community-month) terms. To show the gains from inclusion of the pretreatment covariate, we estimate two versions of the main regression specification, as follows:

$$Y_{i,m,2018} = \beta_0 + \beta_1 M_{i,2018} + \beta_2 R_{i,2018} \qquad \text{No covariate adjustment}$$

$$Y_{i,m,2018} = \beta_0 + \beta_1 M_{i,2018} + \beta_2 R_{i,2018} + \gamma \bar{Y}_{i,m,2012-2017} \qquad \text{Adjustment for 6-year average by month}$$

We varying the effect size for the main monitoring treatment (common arm) and the state engagement arm, we show power estimates for both one-tailed tests for the estimator β_1 in the preceding equations in figure Table 9. We cluster standard errors at the community level. This simulation invokes strong assumptions as to the year-to-year seasonal variation. Our assumptions about seasonality and patterns of deforestation underpinning this simulation are drawn roughly from existing data. The gains from covariate adjustment will rely heavily on the correlation between the lagged dependent variable and realized outcomes. In the simulations, the correlation between potential outcomes (unit-month level) in two given years is $\approx$.5. Increasing this correlation leads to substantial gains in power while reducing it leads to losses in power with the covariate-adjusted estimators. Considering our analyses of pre-treatment data, we anticipate that we can detect a standardized 0.18 effect with a power of 0.8 with the covariate-adjusted estimator, as suggested by Table 9.

	Simulated Standardized ATE of Monitoring			
Estimator	0	0.1	0.2	0.3
No Adjustment	0.048	0.114	0.276	0.46
Adjustment	0.052	0.400	0.840	0.984

Table 9: Power for one-tailed tests (left and right, respectively) using simulated survey data without covariate adjustment. Each cell represents the estimate over 500 simulated data sets.

10 Attrition and Missing Data

Given that our outcome data is based on publicly available satellite data, there is no risk of attrition *per se*. While satellite observations are sometimes compromised by cloud cover or other physical phenomena, such features are uncorrelated with treatment assignment and have been discussed bin detail in the exposition of this data.

The greatest threat of attrition with the survey data comes from a community refusing to participate in the endline survey. RFUS will endeavor to mitigate such problems by identifying uninterested communities in the community selection process and defining the experimental sample accordingly.

11 Contribution to the Meta-analysis

The meta-analysis associated with this Metaketa is preregistered with EGAP. To the extent that the tests and estimators depart from those pre-specified for the meta-analysis, our contribution to the meta-analysis will adhere to the specifications in that pre-analysis plan.[9]

12 Risk Management Plan

12.1 Weather setbacks (extremely dry or wet)

The Amazon region of Peru still maintains vast expanses of forested land with little presence of state institutions and infrastructure. While this makes an ideal context for the monitoring intervention, it also complicates travel and project logistics. All intervention communities are located along major rivers which are a principal means of ingress and egress. Therefore, weather conditions, including extreme dry conditions during the summer and heavy rains and flooding

[9] See the Meta-Analysis Pre-Analysis Plan at `http://egap.org/registration/2815`.

during the winter months, can cause delay in travel between city centers (e.g. Iquitos) and intervention communities. Heavy rains during winter months can limit the ability to travel by road and by foot due to poor roads or trail conditions. These conditions will not adversely affect the project as the primary mode of travel to communities is river. However, during the summer, long dry periods can reduce river levels cause travel times by boat to increase by 1-2 days.

Ultimately, there is little we can do to prevent or predict delays caused by weather. However, our project timeline takes into account seasonal weather conditions and our design mitigates the risk of extreme weather undoing the randomization process (for example, through differentiated treatment to communities due to travel distance). First, the principal risk due to weather is from extremely dry conditions limiting travel to communities during the summer months. This could cause delays in the implementation of the treatment (monitor training) and in the follow up travel to treatment communities. We mitigate the first of these problems by holding trainings during the beginning of the wet season, when travel conditions to communities are better. The second problem (follow up trips), is likely a lower risk. These trips are done once a month per river basin, and all communities will be visited during the same trip. To avoid differentiated followup for treatment communities, we will begin the trip with the furthest to reach community, and work backward toward the urban center and the easier to reach communities. Therefore, the followup trips will only be conducted if the most difficult to reach community is accesible. If travel to the most distant community is impossible during a given week, then all communities within the basin will have a delayed followup trip.

12.2 Political risks with indigenous populations

There are certain political risks inherent to working with indigenous communities in the Peruvian Amazon because of the potential for difficulties when coordinating with indigenous federations. Each community belongs to a different regional federation, which organizes villages into larger political units based on geographic and sometime ethnic divisions. These federations then serve as de facto regional and national governments uniting diverse communities. Although there is no legal obligation to coordinate with these federations, practically and ethically speaking, having sound relations with these federations is essential for a successful project.

However, coordination with federation leaders also brings additional risks to the project. First, leadership can often change due to periodic elections (which are not always held on a regular basis) or due to corruption scandals. In addition, a corrupt leader could extort the project by threatening to impede access to communities unless project money is distributed directly through leaders. Finally, a federation leader could try to interfere with the randomization process to funnel project funds through preferred communities that support the leader.

The principal means through which we will attempt to mitigate these risks is through the knowledge and experience of our implementing partner RFUS. Rainforest has two local staff in charge of project implementation that have many years of experience coordinating and working within indigenous federations in Peru. They have been assessing the political climate of the different regional federations that have jurisdiction over potential project sites, and are using these criteria for making the final selection of river basins. RFUS has suggested that the greatest chance of success for the project will be to work with the ORPIO federation, which has jurisdiction over two potential project sites, the Amazon and the Marañón River Basins. ORPIO?s current leadership has a stable base of support, has demonstrated good management practices with funding from other projects, and has a good relationship with RFUS staff. RFUS has also suggested delaying implementation of the project until after June federation elections, a decision that we have already incorporated into the project timeline.

12.3 Community level - people dropping out of study or cause conflict

There are also risks at the community level that may affect the project and cause non-compliance. We have identified two potential areas of difficulty. First, a community may ultimately not be interested in participating in the monitoring intervention or may be plagued with internal conflict that would inhibit participation. Second, communities that desire to participate and are selected for the control condition may be resentful, and demand to be compensated with additional project funds or create a monitoring project of their own. The first problem would create a situation of one-sided non-compliance (never-takers), while the second would imply two-sided non-compliance (always-wakers).

To mitigate the first problem, we conducted a pre-selection process in which we require all the communities to manifest an initial willingness to participate in the intervention. RFUS will first hold informational sessions in which representatives of all communities in a given river basin will participate. Each community that desires to participate will then discuss the possibility internally within their community and send RFUS a singed statement indicating their interest in participating. The randomization will then be conducted on the basis of this final list of pre-selected communities.

For the second non-compliance problem, RFUS has advised that the best way of dealing with resentful communities is through an open and transparent selection process. RFUS will emphasize with communities that the selection will be through a lottery and completely fair. RFUS will also have a continued presence in each of the project areas, and has already expressed that it will continue to implement the project in these basins if it proves successful. Therefore, RFUS will also be able to offer the intervention to control communities at a future time, which will reduce the risk of resentment by control communities. In addition, given the training and access to technology involved in the monitoring project, it will be difficult for control communities to implement an intervention without the help of an outside NGO or foundation.

12.4 Illegal occupation of land for coca and gold extraction

There is a constant risk in the Peruvian Amazon that indigenous land is occupied for means of illegal activities such as coca cultivation, gold mining, or illegal logging. These risks are endemic to the region and affect communities regardless or participation in the intervention. In addition, greater information provided by monitoring teams will potentially allow communities to better limit exposure to these risks and alert national authorities. However, involvement in the monitoring project could expose monitoring teams themselves to a greater danger due to the presence of illegal activity on indigenous land. We will attempt to mitigate this risk for all participants in two ways: (1) by ensuring that all participants are fully aware and adequately trained in risk prevention; and (2) through the continuing involvement and support of our implementing partner, RFUS.

The first method of risk prevention will be incorporated into the training of community monitoring teams. These teams, given their relatively consistent presence in the community?s forest, have the greatest potential to be exposed to elevated levels of risk and therefore will be the focus of our risk prevention efforts. A large segment of the training program for monitors will involve a participatory workshop that will help community members identify and reduce potential risks. The training itself will emphasize conflict prevention and methods for avoiding confrontation with non-community members engaged in illegal activity. However, it is important to note that the community members themselves are in possession of the greatest knowledge of the risks in their lands and how to mitigate these. Therefore, the training will primarily utilize participatory methodology that will help community members apply their knowledge of the area to identify risks and plan how to limit these accordingly.

The second method for mitigating these risks will be through the continued involvement of RFUS, which already has a program in place for assisting communities threatened by illegal loggers or squatters. This program involves funding required medical, legal, or logistical assistance to attend to any individual injured by illegal activity on community lands and to advocate for appropriate state intervention seeking criminal liability for those responsible and adequate compensation for the community. RFUS also has an extensive network of organizations and advocates both regionally in Ucayali and Loreto, and nationally in Peru, that can help provide assistance to all communities participating in the project.

References

Agrawal, Arun. 2007. "Forests, Governance, and Sustainability: Common Property Theory and its Contributions." *International Journal of the Commons* 1 (1): 111–136.

Avelino, Andre Fernandes Tomon, Kathy Baylis, and Jordi Honey-Rosés. 2016. "Goldilocks and the Raster Grid: Selecting Scale when Evaluating Conservation Programs." *PLOS ONE* 11 (12): 1–24.

Benjamini, Yoav, and Yosef Hochberg. 1995. "Controlling the False Discovery Rate: A Practical and Powerful Approach to Mulitple Testing." *Journal of the Royal Statistical Society. Series B.* 57 (1): 289–300.

Casey, Katherine, Rachel Glennerster, and Edward Miguel. 2012. "Reshaping Institutions: Evidence on Aid Impacts Using a Preanalysis Plan." *Quarterly Journal of Economics* 127 (4): 1755–1812.

Dietz, Thomas, Elinor Ostrom, and Paul C Stern. 2003. "The Struggle to Govern the Commons." *Science* 302 (5652): 1907–1912.

IBC. 2016. Directorio 2016 Comunidades Nativas del Peru. Technical report Instituto del Bien Comun.

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

Ostrom, Elinor. 1990. *Governing the Commons*. Cambridge university press.

Ponce, Nataly, Eli Castillo, Liliana Bances, and Cesar Ipenza. 2015. Study on the Criminal Justice System in Environmental Matters in Peru–Executive Report. Executive report USAID.

Yashar, Deborah J. 2005. *Contesting Citizenship in Latin America: The Rise of Indigenous Movements and the Postliberal Challenge*. Cambridge University Press.