Watchman Field Experiment – Pre Analysis Plan

Julian Dyer

February 18, 2019

Abstract

This document describes the analysis plan for a field experiment in crime and property security in rural Kenya. I create randomised variation in farm security by matching Kenyan smallholding farmers with subsidized Maasai watchmen to protect farms. I explore the effect of theft and property crime by testing whether farmers engage in different types of production when their farms are more secure against theft. This project also examines effects on time allocated away from the household, and whether improved security increases interaction with agricultural markets. I further explore how an environment of weak enforcement and imperfect information on the identity of thieves influences social relationships, and whether improved security and reduced fear of crime decrease the degree of ethnic ingroup-outgroup parochialism. Finally, I test for spillover effects and responses by local institutions.

Keywords: theft, crime, institutions, agriculture

Contents

1 Introduction

This paper studies the consequences securing farms in Migori County, Kenya, against theft. In this setting, smallholding farming is the primary economic activity and where formal institutions imperfectly protect property rights. Fears of farm theft are pervasive, and can distort production decisions away from high-risk crops towards less easily stolen staples. This risk is increased when farmers are away from their households, leading to less time spent off-farm and less time interacting with local markets for agricultural products. The difficulty in identifying who is responsible leads to suspicion of others in the village, where other ethnic groups are often blamed for crime. This is also a context where public provision of property rights can readily be replaced by private protection, allowing households to avoid relying on ineffective public institutions. It is, however, difficult to identify causal mechanisms from observational data, as culture, institutional quality and production decisions are determined simultaneously. To understand how theft influences social outcomes and production decisions, I conduct a novel randomised experiment in protection against property crime. I randomly assign private protection of farms through an intervention facilitating the hiring of highly trusted Maasai watchmen, by matching households with watchmen and through targeted subsidies for wages and up-front payment of travel costs.

2 Intervention

The intervention in this project is matching farming households to high-quality, trusted Maasai watchmen at a heavily subsidized rate. The intention of this intervention is to cause variation in the security of farms during the short rains season, beginning with planting in August. Watchmen are recruited with the assistance of partners from the Maasai Education Research and Conservation Centre (MERC) in Maasailand in January and early February. After information sessions with farmers and the collection of supplementary baseline information, farmers will be informed of their treatment status by the field coordinator in early June, giving them time to adjust planting decisions and input purchases. After being informed of their treatment status and given the contact information of the watchman they have been matched to, I schedule times for farmers to be expect a phone call from a watchman, so they can arrange the time of employment and other details. The choice of watchmen as the security intervention for this project was motivated by the fact that many other security interventions (such as fencing) would include a significant element of improved security of land tenure in addition to security from crime and theft. A fencing intervention, for example, would first require demarcation and clarification of exact boundaries and the status of land to be fenced, which in itself would have a strong effect on land tenure which is well known to impact agricultural decisionmaking.

For this study to successfully test whether cropping decisions are influenced by security, it was crucial that farmers were credibly informed of their treatment status. For this reason, the intervention included three separate attempts to inform them. First, farmers received phone calls from Busara Centre staff informing them of their status, and informing treated farmers to expect a call from a watchman. Second, the watchman coordinator ensured that all watchmen called their matched farmer during the assigned time frame. The watchman coordinator also verified that they had successfully communicated with the matched farmer, arranging for interpreters who could translate into local languages where the watchman and farmer struggled to communicate in Swahili. Finally, the local farmer coordinators followed up with farmers after these first two attempts to confirm that they knew their status, and to inform the watchman coordinator if any treated farmers had not yet spoken with their assigned watchman. All three of these rounds of information occurred by early July, allowing a generous amount of time for farmers to consider cropping decisions and adjust their inputs and potentially learn about new crops they might want to plant. The wage rate paid by farmers and the subsidy are set in advance, so the treatment is uniform across the sample. The duration of the treatment is also set at a uniform six weeks of watchman employment, at a time set by farmers to coincide with when they anticipate their crops will be at risk.

A potential risk for the success of this intervention was that the Maasai watchmen would feel uncomfortable being in a new area and choose to withdraw from the study. To avoid issues, three additional Maasai coordinators were deployed to Migori a week prior to the first deployment of watchmen to farms, to prepare the farmers, greet the watchmen as they arrived and direct them to reach their assigned farming households, as well as working with local farmer coordinators to ensure they were able to find accommodation. These Maasai coordinators remained in Migori for the duration of the study, to help watchmen with any minor issues that arose and to check that the watchmen were strictly being asked by the farmers to do security work to ensure that the intervention did not unintentionally provide subsidised farm labour. Farmers were also informed during recruitment that the watchmen would strictly do security work, so they would not have been anticipating additional farm labour while making planting decisions.

To ensure all farmers in the sample have the same access to information and crop inputs, all farmers in the sample receive the exact same notification by SMS of information and training sessions hosted by the local KALRO affiliate organization, Community Action for Rural Development (CARD). These extension programs are the exact same program currently offered by CARD in this region and is accessible to all farmers in the region, and is typical of the agricultural extension services that are prevalent in rural areas of less-developed countries in general. Access and invitation to these sessions is completely even across treatment and control groups.

2.1 Theory of Change

Imperfect property security and the fear of theft constrain economic activity by increasing the risk to specific types of economic activity. Two examples of this are the planting of high-value 'stealable' crops and taking work or starting enterprises away from the farm. Non-experimental evidence suggests that some types of agriculture are constrained by perceived theft risk. (Schechter 2007) Theft risk is perceived to be higher for crops that are high-value per kilogram, are easily picked, are sold on open markets (as opposed to crops like tobacco and sugarcane which are sold on closed markets to cooperatives or processing plants), and have a long window of maturity. In addition, crops that are grown by fewer farmers are also higher risk, as new/different crops are perceived to make a farmer a target for theft. Similarly, spending time away from the farm is perceived to increase the risk of theft of crops during the farming season as farms are perceived to be insecure. In this way, imperfect security acts to discourage farmers from engaging in off-farm economic activities.

I therefore hypothesise that an intervention to improve security of farms will reduce the risk to these types of economic activity, and will lead to increased land allocated to high theft-risk agriculture and greater likelihood of pursuing economic activities away from the farm.

In particular, this will impact two aspects of off-farm economic activity. Farmers report concerns about leaving their farm regularly, as they believed people would notice when they were away, and that their farms would be targeted if others in the area realised they weren't home. As a result, this constraints their ability to spend time on off-farm enterprises (like small businesses, market stalls, etc.). As a result, this intervention will allow farmers to spend more time away from their farms, on economic activity and also on non-economic activity. This intervention will also have an effect on the marketing of crops through two channels. As high-value crops are most at risk of theft, this intervention to improve security will lead to farmers growing more crops that are high-value and thus worth marketing off-farm. In addition, this security intervention will reduce the risk to farmers from regularly leaving their farm to go to the market to sell crops, where they receive a higher price than when they sell at the farm gate to middlemen. Therefore, by reducing the risk to both producing high-value crops, and to leaving the farm to market those crops, this intervention will cause farmers to sell a greater share of their crops away from their farm, spend more time on marketing activities and receive higher prices for their sales.

There are a number of other interesting questions that I consider in this project. These are hypotheses where I have relatively weaker priors, these are more exploratory outcomes. The nature, or more accurately the *perceived* nature, of property insecurity is of great importance for understanding whether interventions in property security have broad spillovers. The existence of positive externalities would justify increased spending on security and state-building at the micro-level as valuable public goods, where private decisions focused only on direct personal benefits lead to sub-optimal investment.

I consider two broad types of spillovers. The first are direct benefits to the effectiveness of coordinating investments in security. One perceived explanation for crime in villages is that there is a fixed number of thieves that cause property risk. If this is the case, then as the number of farmers pursuing higher-risk strategies (either planting more valuable crops or pursuing more off-farm economic activities) increases then there are more potential targets for theft, and the risk to any individual pursuing higher-risk activities is lower.

The other type of positive externalities are potential indirect benefits to others through social learning, improved social cohesion and effects on legitimacy and trust in institutions. If improved property security leads to greater experimentation by farmers, the knowledge gained from this experimentation may spread to other farmers in their community. This could be direct knowledge about agricultural production, or it could be information about the actual risk of crime. As expectations of crime risk are mostly hypothetical, and not learned through own experience, then it is possible that the risk of crime is over-estimated. If farmers experiment with high-risk activities and find their security is less active and they face fewer theft attempts than expected, they may revise their expected frequency of theft events downwards, and this knowledge may also diffuse to other farmers. If it is indeed true that *'good fences make good neighbours'* then improved farm security may reduce disputes. Farm security may also lead to broader social networks (if it causes greater off-farm employment or gives greater incentives for intra-village trade through increased specialization) and less ethnic parochialism, if out-groups are stereotyped as being responsible for crime.

2.1.1 Logic of Argument

An intervention to improve security relaxes constraints on crop choices and time spent on off-farm economic activities.

i. Agricultural Production:

   (a) Subsidized farm watchmen improve subjective farm security and reduces risk of theft from planting valuable crops

   (b) With subsidized protection, farmers change their crop decisions, planting more theft-prone and higher value crops

   (c) With reduced risk of theft, farmers more willing to experiment with new crops. From growing these new crops for the first time, they learn about these new technologies and having already borne the learning costs they continue to grow these crops in the future.

ii. Off-Farm Enterprises & Employment

   (a) Subsidized farm watchmen improve subjective farm security and reduces risk of leaving farm unattended to do work away from the farm

   (b) Improved security should lead to more farmers taking work off-farm and starting more off-farm enterprises.

iii. Market Interaction & Prices

   (a) Reduced risk of theft from planting high-value crops leads to increased production of crops with ready markets away from harvest cycle of main staples.

      i. This increases the potential returns from spending time selling crops away from the farm.

   (b) Improved security means reduced concerns from regularly spending time away from the farm.

      i. This reduces the potential risk of spending time marketing crops.

   (c) By these two channels, farmers will spend more time selling their crops away from their farm, will sell a greater share of their crops off-farm. As a result, they will receive higher prices than when they sell to middlemen who buy crops at the farm gate.

3 Research Questions

The main research questions for this paper concern the impact of insecurity and fear of crime on economic activity in rural communities.

The main hypothesis for this paper and exploratory hypotheses are outlined below.

3.1 Main Hypotheses

The primary hypothesis to be tested is that perceived farm insecurity restricts economic activity. Specifically, it influences economic activity in the following three ways:

i. **Crop Choice.** Imperfect property security constrains the type of crops farmers choose to plant. An intervention to improve perceived farm security will allow farmets to plant new theft-risky crops and allocate more land to theft-prone crops they were already growing.

ii. **Time Off-Farm.** Fears about theft and interference on the farm while a farmer is away restricts the amount of time they spend off-farm on economic or non-economic activities.

iii. **Market Interaction.** Improved property security allows farmers to shift production towards more crops for sale in the local market, and allows farmers to spend more time marketing their crops off-farm. This leads to improved prices for their crops and a greater share of their production being sold off-farm.

3.2 Secondary Hypotheses

I expect that farm security will influence a number of other important aspects of behaviour, but as these outcomes occur as a result of the changes described in the primary outcomes, these are treated as secondary outcomes.

i. **Household Profits.** As a result of the changes to economic behaviour outlined in the above main hypotheses, household profits (the sum of farm profits, employment earnings and earnings from other enterprises) should increase.

ii. **Learning.** A temporary improvement in security causes changes in economic production which are persistent. Specifically, farmers learn about new crops they intend to keep planting or learn about the benefits of time spent off-farm and as a result their future willingness to pay for a watchman increases.

iii. **Behavioural Updating.** A temporary improvement in security causes persistent changes to attitudes in ways that are not rational. Specifically, I look for effects on:

   - **Expectations of theft next season.** I explore whether farmers display some degree of attribution bias, and incorrectly update their expectations of theft, reducing their suspicion of neighbours.
   - **Ethnic attitudes and stereotyping.** In a context where ethnic outgroups are often blamed for theft and crime, if the security intervention reduces ethnic stereotyping this would be evidence of scapegoating as a form of motivated belief that occurs in situations of vulnerability.
   - **Attitudes towards formal institutions.** Respondents report at baseline that they think their chiefs are doing a good job enforcing property rights and protecting from crime, while also not expecting they will be able to do anything to punish thieves. If the security intervention causes people to lower their opinion of the chiefs, then this is evidence that institutions compete for legitimacy, and that the existence of a private option weakens the legitimacy of the formal.

iv. **Neighbour Disputes & Grievances.** A context of insecure property, ineffective enforcement institutions and imperfect information on identity of thieves causes disputes among neighbours. In addition, the high social costs to reporting someone mean that many grievances aren't addressed and manifest as resentment. The property security intervention will reduce disputes and grievances among neighbours.

v. **Crime Spillovers.** Farm theft is perceived to be local and opportunistic, so an intervention to improve security in one part of a village will not have significant displacement effects on crime experienced by farmers elsewhere in the village.

vi. **Agricultural Learning.** It is well known that there is a significant amount of social learning in agriculture. An intervention that allows farmers to experiment with new crops will therefore lead to greater knowledge spillovers within the village. I test whether neighbours report greater learning about new crops or gain new information on marketing in treated villages. If these positive spillovers are strong enough, this would suggest that security interventions are worth subsidizing, even if individual WTP is not high enough to justify the cost.

vii. **Institutional Substitution.** When formal institutions are replaced by a more effective non-state alternative, actors in the formal institutions increase their visible activity to protect their legitimacy.

viii. **Social Networks.** If in the main outcomes I find that farmers are spending more time away from their farms, I will then test the downstream effects on their social interactions. The outcomes of interest here are whether they built new business relationships, and whether they were more likely to have interacted and cooperated with people outside their usual network. If farmers build broader networks then this suggests that insecurity causes parochialism through an interactions channel, where the expanded economic activity allowed by security causes farmers to broaden their social horizons.

ix. **Demand for Agricultural Extension.** Using CARD administrative information on attendance at different training sessions, I can test the hypothesis that treated farmers are more likely to attend training sessions to learn about new crops.

3.3 Main Outcome Variables - Definition

Here I describe the outcome variables used to test my main hypotheses, including the manipulation checks to establish the mechanism. Only the indices for the main outcomes, indicated with † will be corrected for multiple testing.

i. **Manipulation Check**
   (a) Intervention Take-Up
      - Extensive: Farm had a watchman (1/0)
      - Intensive: How many weeks they worked
      - Confidence: Were they confident the watchman was working hard (1/)

   (b) Perceived Security Index
      - Low farm security this season (1/0)
      - High perceived theft risk: high-value crops (1/0)
      - High perceived theft risk: different crops (1/0)
      - Perceived likelihood of theft attempts

ii. **Crop Choice Index**†
   (a) Total change in land use
   (b) Share of land allocated to new crops
   (c) Average Theft Risk
   (d) Land used for theft-risky crops
   (e) Land used for highly theft-risky crops
   (f) Land Share new or increased land crops due to improved security
   (g) Number of new or increased land crops or grown on increased land due to improved security

iii. **Off-Farm Activity Index**†
   (a) Time use share off-farm
   (b) Any off-farm enterprise (1/0)
   (c) Any off-farm employment (1/0)
   (d) Increased time spent off farm (1/0)

iv. **Off-Farm Market Interaction Index**†
   (a) Share of crops sold off farm
   (b) Share of sales taking place off farm
   (c) Increased sales off farm (1/0)
   (d) Average price premium over crop mean

3.4 Outcome Variables - Construction

Here I describe exactly how the outcome variables above will be created from the data collected and the survey questions. I also describe how these outcomes will be combined into a single index for each hypothesis.

i. **Manipulation Check**

   (a) Intervention Take-Up

      i. Extensive Margin: "Did you hire a watchman in the past Short Rains season (beginning in August-September)?"

      ii. Intensive Margin: "For how many weeks was this watchman protecting your farm?"

      iii. Perceived Effort: "Are you confident your watchman was working hard guarding your farm?"

   (b) Perceived Security Index, composed of the following:

      i. "How well protected was your farm this season?" (reported 4 or 5)

      ii. "If in this last short rains season you planted high-value crops, how likely is it that they would have been stolen?" (reported 4 or 5)

      iii. "If in this last short rains season you planted different crops from others around you, how likely is it that they would have been stolen?" (reported 4 or 5)

      iv. "How likely was it that someone near your farm would try and steal from you this season?"

ii. **Crop Choice Index**[†]

   (a) Total change in land use

      i. Using baseline and endline data on land use by crops, take the total amount of land use change:

$$\text{LandUseChange}_i = \sum_{c=1}^{40} |\Delta l_{c,i}|$$

   (b) Share of land allocated to new crops

      i. Constructed from survey data on land allocated to different crops, combined with questions asking if this was the first time they have grown this crop. The variable $\text{LandshareNewCrops}_i$ is the share of land allocated to new crops the farmer has never grown before.

   (c) Land used for high theft-risk crops

      i. I construct a theft-risk score for crops as an index of objective crop characteristics qualitative evidence suggests is related to theft risk: price per kg, time to harvest per kg, open/closed markets, consumed locally and length of harvest window. The final characteristic is the share of farmers growing the crop at baseline, as farmers perceive that different or uncommon crops are more at risk. I take the sum of these variables, after standardizing, and use the sum as a theft-risk score. I then use three different measures of land allocated to high theft-risk crops. The first is average theft riskiness of crop choice constructed as follows

$$\text{AvgTheftRisk}_i = \sum_{c=1}^{40} \text{LandShare}_{c,i} \cdot \text{TheftRisk}_c$$

      where i indexes individual farmers and c indexes crops.

ii. The second measure is constructed by defining all crops with above-median theft risk scores as being high risk, then calculating the share of land allocated to high-risk crops under this definition.

iii. The third measure is similar, but defining high-risk crops as being those with above 75th percentile theft risk and calculating the share of land allocated to these crops. These three measures of theft-riskiness of crops will be combined into a single index.

(d) Share of Land used for crops that were either new or had their land allocation increased as a result of improved security. When a farmer reported growing a new crop or increasing their land allocation to this crop, they were offered multiple choice list of reasons, one of which was reduced fear of theft/interference on their farm. If this option was selected, the crop is counted as a crop who was new or had increased allocation due to security.

(e) Number of new crops grown or grown on increased land due to improved security. New crops or crops with increased production due to security are defined as above.

iii. **Off-Farm Activity Index**†

(a) Time use share off-farm

i. This measure is constructed using the responses to *"In the last month before harvest, how much time did you spend (in minutes) during an average day on ..."* with the off-farm time use categories being Marketing crops (off-farm), leisure time (away from home),other economic activity (away from home) and non-leisure social activity (faith based, community groups, etc.) off farm. The on-farm categories are tending to crops, leisure time (at home), other economic activity (at home), and household chores.

(b) Any off-farm enterprise (1/0)

(c) Any off-farm employment (1/0)

(d) Increased time spent off farm (1/0)

i. This is a self-reported question, where the farmer was asked if they spend the same amount of time, more time or less time off the farm than the same season last year. If they answered more time off farm, this is coded as 1, otherwise 0.

iv. **Off-Farm Market Interaction Index**†

(a) Share of Crops sold Off-Farm

i. Constructed as the share of crops that are sold off-farm at all. Any crops where the respondent selected one of the off-farm options is coded as being sold off-farm.

(b) Share of sales off-farm

i. This is constructed by, for each crop, taking the share of sales that were away from the farm. I then take the average across crops with nonzero sales.

(c) Avg Price deviation (%) above crop mean

i. To construct this, I first compute the weighted average price for each crop for each farmer, where the prices received in different markets are weighted by the share of sales sold in that market. For each crop, I take the cross-farmer average of these weighted prices. Using this cross-farmer average weighted price, I then compute for each farmer their deviation from this average price, where i indexes individuals and c indexes crops.:

$$\text{PriceDeviation}_{c,i} = \frac{\text{WeightedPrice}_{c,i} - \text{AvgWeightedPrice}_c}{\text{AvgWeightedPrice}_c} \tag{1}$$

For each farmer, I then take the average of their $\text{PriceDeviation}_{c,i}$ across crops.

(d) Spent more time marketing crops off-farm during this season this year than during the same season last year. (1/0)

4 Experimental Design

4.1 Sampling and Recruitment

This experiment consists of a core sample, answering the main research questions, and a supplementary spillover sample that will be used to estimate the spillovers of a crime intervention.

4.1.1 Main Sample

The main sample of farmers for this experiment are drawn from the field networks of the Kenyan Agricultural and Livestock Research Organization (KALRO) in Migori county. The local KALRO affiliate in Migori County is the organization Community Action for Rural Development (CARD) who maintains connections with farmers through a grassroots Farmer Research Network (FRN) which empowers farmers to undertake grassroots research projects where the community chooses research topics. This region was selected for lack of ethnic hostility towards Maasai as well as proximity to Maasailand, meaning transport is feasible. Migori was not selected for it's agricultural potential, and the conditions in the region are roughly typical of Kenya. The agricultural conditions in Migori allow for planting of some horticultural crops in addition to local staples, and the selected sub-counties are a reasonable distance from Migori town and other urban centres, giving an opportunity for farmers to seek off-farm employment during this planting season.

A target sample of ten farmers was selected in sixty villages for a total of 600 farmers in the core sample. This sample was recruited using the farmer networks maintained by the Kenya Agriculture and Livestock Research Organization (KALRO). This recruitment procedure was designed to mimic the standard mobilization procedures used by KALRO in their regular agricultural extension programming. After sixty villages in three sub-counties (Suna East, Suna West and Uriri) near Migori town were identified, information meetings explaining the project and intervention were conducted with leadership of the farmer's group in each village.[1] Ten farmers were selected at each meeting within the village's farmers' group, who were then invited to a session where they signed consent forms and baseline data was collected. Some logistical issues arose which impacted turnout from some villages at the consent sessions, such as clashes with a local market day or funeral.[2] Respondents from some villages whose representatives had not been present mobilized recruits for the consent sessions in their area, so they were included in the sample to increase the number of clusters and improve

[1] See Appendix for information session script.

[2] The ability to respond to these issues and mobilize alternate respondents varied with the quality of local coordinators. My local partner was uncomfortable with over-inviting people to information sessions given the cost and inconvenience of coming to sessions, and the cost of potentially paying many extra respondents was a concern. A particularly prescient concern was potential resentment from invited respondents whose villages were assigned to treatment but who were not included and were not matched with a subsidized watchman.

statistical power. The final sample of respondents was 599 respondents in 76 villages. The consent and survey sessions with individual farmers take place from May 29th to June 6th.

4.1.2 Spillover Sample

In order to estimate the spillover effects of the intervention on the research questions, a supplementary spillover sample will be recruited. From the original sample of small geographic clusters, twenty treated clusters and ten control clusters will be selected by stratified randomisation. In each of those clusters, enumerators will be sent into the villages to record the GPS locations of the treated households. Enumerators will then conduct a convenience sample of other farming households in the village at varying distances to the households in the main sample. In total the spillover sample will include three hundred households.

4.2 Randomization Units & Method

The unit of randomization will be the village. This is motivated by the potential for interaction between treated and control, and to ensure that at least two watchmen are assigned to each village and that no watchman has to travel alone to a new place.

Clusters will be assigned to treatment using stratified randomization.[3] Following Athey & Imbens, the ideal is to stratify as much as possible while ensuring that there are at least two treatment and two control units, so that variances can be estimated. In my case, where the share treatment is two-fifths, this means that to have two treated in each stratum requires five village units per stratum. Within this constraint of five villages per stratum, I plan to stratify on the following variables:

i. **Sub-County**. I first stratify by the three sub-counties (Suna East, Suna West and Uriri) included in the sample. This is to control for geographic endowments and the type of agriculture practiced, along with other unobserved regional differences.

ii. **Village Number of Respondents**. As discussed above, there was variation in number of respondents across villages, so stratifying by number of respondents helps to ensure that the variation in cluster size does not lead to imbalance at the individual level.

iii. **Village Mean WTP For Watchman**. Self-reported willingness and ability to pay for a watchman acts as an aggregate measure of how important security is in a particular village and farmer income, so is an important variable for stratification.

Stratifying by these variables partitions will divide the sample into 12 strata. This stratification will be implemented and deal with misfits by the within-stratum approach, as described by Carril (2017) and Bruhn and McKenzie (2011).[4]

[3] Random Seed: 69052039. Generated from https://www.random.org/ on May 19, 2018.

[4] The choice of the within-stratum was motivated by the fact that the treatment proportions and number of village units will generate a reasonable number of misfits, so I chose the method which preserved balance at the cost of harming treatment fractions: "*wstrata* randomly allocates misfits independently across all strata, weighting treatments as specified in unequal. This ensures that the fractions specified in unequal() affect the within-distribution of treatments among misfits, so overall balance of unequal treatments should be (almost) attained. However, this method doesn't ensure the balance of misfits' treatment allocation within each stratum (they could differ by more than 1)." Carril (2017)

4.3 Characteristics & Power

Given the sample size of 600 in approximately 60 clusters of 10 respondents each, the power calculations for my main outcomes, relating to the choice to plant theft-risky crops, are as follows. This design will be able to detect a 4% increase, corresponding to 0.26 standard deviations, in land allocated to high theft-risk crops relative to the district-level average, where high theft-risk crops are designated using objective characteristics of crops.

5 Data

5.1 Data Collection Methods and Instruments

5.1.1 Data Collection

Data collection was conducted using the SurveyCTO platform on Android tablets used by survey enumerators trained by Busara Centre staff. Baseline data was collected prior to assignment to watchmen treatment as exit surveys after the consent & information sessions.[5] Respondents came to central locations in each of the three study sub-counties where the baseline surveys were conducted privately by trained and experienced enumerators. Endline data was collected by household visits using local guides and farmer coordinators to locate sample households. Backchecks were implemented for a subset of this sample to check the accuracy of the data.

The survey included a number of other measures to ensure data quality. There were a number of built-in checks for inconsistent data that notified the enumerator during the survey of any issues (such as the sum of harvest sold at different markets not being roughly equal to the reported total amount sold, etc.) In addition, the survey included time checks where it was recorded if the enumerator took less than a minimum amount of time to complete a section. In addition to the phone backchecks, these features allow for high-frequency checking of data each day during data collection. The Busara coordinator moved around the village while data was being collected to drop in on enumerators and confirm that they were in fact moving through the households they had been assigned. While the supervisor travelled between households to monitor enumerators, they will stop to sample neighbours and other households at varying distances to the main sample households.

5.1.2 Data Management

Identifiable survey data will be dealt with under the standard protocol used by the Busara Centre when conducting surveys, which is as follows. All respondents will be given a unique ID code, and personal identifiers will be saved in separate files from the survey responses. Survey response data collected using the Open Design Kit (ODK) platform is first stored on tablets used by enumerators. All data will be encrypted on the tablets upon completion of the survey and transferred through a secure connection to computers used by Busara staff, where it is automatically compiled onto the ODK server.

To keep data secure, all the netbooks and computers used by field staff will password protected; the passwords are managed by senior staff, are regularly changed, and security is evaluated periodically.

Spillover data will be collected for randomly selected villages by the Busara supervisor in the villages. Enumerators will visit all households in the main sample in the given villages to record their position by GPS.

[5] See Appendix for full supplementary survey instrument.

5.1.3 Demand for Agricultural Extension

As mentioned above, CARD offers training sessions open to farmers in this area to learn techniques for planting crops and get information about accessing inputs. All farmers in this experiment will receive the same SMS invites to these sessions, and attendance will be taken at sessions and later matched to watchman treatment status. This data provides a secondary way to test if improved security increases the demand for information on new crops and improved inputs, and which crops they become interested in.

5.2 Risk and Treatment of Attrition

Attrition is not a significant concern in this study due to the good relationship between the KALRO field coordinator and the farmers. Each village had a local CARD coordinator who maintains close relationships with farmers, which improves our ability to communicate with and mobilize them for surveying. In addition, this study takes place over a single agricultural season, so the timeline is reasonably short and it is unlikely they will have moved or changed their phone numbers by the time of the endline survey, immediately after the harvest. By collecting endline data at the farm households, we reduce the risk of not being able to contact farmers, or them being unable to travel to a central location.

The timing of the endline survey was carefully chosen to coincide with the period between harvesting in early February 2019 and planting in mid-March, so that we were trying to survey farmers during a window where they have the maximum amount of available time to participate.

5.3 Crop Theft-Risk Data

I will use a number of data sources to classify crops as being theft-risky or not. First, high theft-risk is defined subjectively as being the crops mentioned in qualitative interviews as being theft risk. This will be supplemented with an expert survey of a small sample of local agricultural experts. This will be supplemented by an objective measure of theft-risk based on the following categories that qualitative evidence suggest are important for theft-riskiness:

- Price per KG (reflective of resale value, and also local preferences over theft for consumption)
- Time to harvest (easy-to-pick crops are perceived as being a bigger risk than those where a great deal of labour is required to harvest)
- Sold in open local markets or consumed locally (crops such as unprocessed tea or tobacco are only sold to processing plants, not consumed locally, and are perceived as low-risk for theft)
- Time of harvest relative to most common local staple (the incentive to steal is higher for crops that are mature when food is more scarce prior to the harvest of the main staple)
- Length of Maturity Window (crops that mature all at the same time such as avocadoes or sorghum are perceived to be higher risk than bananas or passionfruit, which mature at staggered times)
- Baseline share of farmers growing this crop (risk to each individual farmer is perceived to decrease in number of farmers growing that crop)

6 Empirical Analysis

6.1 Statistical Methods

6.1.1 Handling Outliers

Where possible, outliers will be handled by imputation. Where this is not possible for inputs, like land allocation by crop or input spending per acre, these will be Winsorized at the 2.5% level. Results will also be reported for the 5% level, but 2.5% will be the level of Winsorizing for the main outcomes. Given the variance in total land, I will Winsor total land owned, rented and farmed at 5%.

For outliers in agricultural yield per acre, or price per kg of harvest, values for outliers will be imputed as a function of primitives. Specifically, for the top & bottom 2.5% of output her acre and price per kg, I will impute values as follows by crop:

i. Exclude observations where input spending per acre is an outlier at 2.5% level, and exclude observations where the variable of interest (either output her acre or price per kg) is an outlier at the 2.5% level.

ii. Estimate the following model:

$$y_{i,c} = \beta_{0,c} + \beta_{1,c} \cdot \text{InputsPerAcre}_{i,c} + \epsilon_{i,c}$$

where i indexes individual farmers and c indexes crop.

iii. Use estimated parameters ($\hat{\beta}_{0,c}$ and $\hat{\beta}_{1,c}$) values for outliers in the variable of interest (either output her acre or price per kg) using Winsorized inputs per acre:

$$\hat{y}_{i,c} = \hat{\beta}_{0,c} + \hat{\beta}_{1,c} \cdot \text{InputsPerAcre}_{i,c}$$

This process will also be repeated at the 5% level, but primary results will be with outliers cleaned at 2.5% level.

6.2 Multiple Hypothesis Testing

To account for multiple hypothesis testing, I will be reporting Family-Wise Error Rate (FWER) and False Discovery Rate (FDR) p-values in addition to naive p-values.

To avoid overly harsh punishment for testing multiple hypotheses, I will be testing a single constructed outcome variable (see Section 3.3) for each of the two main hypotheses outlined in Section 3.1. Testing for the manipulation checks will not be corrected for multiple testing.

6.3 Heterogeneous Effects

There is good reason to expect that the effects of improved security will be more important for widows or other female-headed households where there is no adult male in the home. For this reason, I will explore heterogeneous effects of treatment on the main outcome indices by gender of head of household. As this is not the primary goal of the paper, I will not include this in multiple hypothesis correction.

I will also consider heterogeneous effects on crop choice by distance from the main home to the agricultural plot in question. This is an interesting potential source of heterogeneity, as security concerns are likely more pronounced farther from the home. Specifically, I will be splitting plots by above below the median for distance to house, measured in estimated walking time. I will then run the main treatment specification on the Crop Choice Index and components constructed using the subsamples of 'near' plots and 'far' plots, and I expect that the effect of security will be strongest when looking at the further plots. This suggests that security issues are especially important where land tends to be held in a number of small plots, and suggests a benefit to more active land markets where farmers may consolidate their plots.

6.4 Statistical Model

6.4.1 ITT Specification

My analysis will focus on Intention-to-Treat (ITT) estimates to identify the effect of a farm security intervention on economic decisionmaking.

The empirical specification I will use in this project is straightforward. For outcomes where I have baseline data, the specification will be to have all outcomes Y_i, baseline and endline, as observations and to run the following specification:

$$Y_{i,t,v} = \beta_0 + \beta_1 \text{Treat}_{i,v} + \beta_2 \text{Endline}_t + \beta_3 \text{Treat}_{i,v} \cdot \text{Endline}_t + \mathbf{\Gamma}_i + \epsilon_{i,v}$$

where i indexes individual farmers, t indexes period (endline or baseline) and β_3, the coefficient interaction between Treatment and Endline, is the parameter of interest. All specifications include a vector of fixed effects for randomization strata, $\mathbf{\Gamma}_i$.

6.4.2 Standard Errors

Standard Errors for this project will be clustered at the village level, as treatment is assigned at the village level. I will use standard methods for calculating clustered standard errors, correcting for Multiple Hypothesis Testing by reporting FWER and FDR p-values in addition to naive p-values for my main outcome indices.

7 Administration

7.1 Ethics Approval

This project has been approved by the University of Toronto Research Ethics Board, Protocol #34160.

7.2 Funding

This project has received funding from the International Development Research Centre (IDRC) under a Doctoral Research Award.

This project has also received funding from the JPAL Crime and Violence Initiative (CVI).

7.3 Acknowledgements

Arthur Blouin, Gustavo Bobonis and Marco Gonzalez-Navarro have contributed greatly to this project and provided excellent supervision.

This project has benefited from the help of Laura Derksen, Chaning Jang and Jeremy Shapiro.

References

Bruhn, Miriam and David McKenzie (2011). *Tools of the Trade: Doing Stratified Randomization with Uneven Numbers in Some Strata*. The World Bank: Impact Evaluations. URL: http://blogs.worldbank.org/impactevaluations/tools-of-the-trade-doing-stratified-randomization-with-unequal-numbers-in-some-strata..

Carril, A. (2017). "Dealing with misfits in random treatment assignment". In: *Stata Journal* 17.3, pp. 652–667.

Schechter, Laura (2007). "Theft, Gift-Giving, and Trustworthiness: Honesty Is Its Own Reward in Rural Paraguay". In: *The American Economic Review* 97.5, pp. 1560–1582.

A Recruitment Information Session Script

Hello, my name is Julian Dyer, I'm a PhD student at the University of Toronto and I'm are here to tell you about a research project that will be running in your area. This project is an experiment to understand how security concerns influence your business as farmers.

The way this works is we will empower some farmers with improved security by matching you with high-quality Maasai watchmen from Narok District at a subsidized wage rate. This design has been chosen because of their reputation for being very trustworthy guards and because as outsiders in the community, they will not have other obligations to allow others to steal the crops they are guarding.

Their pay comes mostly from the research budget I have been allocated for this project, but farmers will pay 250 Kenya Shillings per week to the watchmen. The watchmen will be available for up to six weeks to work for the farmer they have been matched with, or less if the farmer chooses they do not want to hire them for that long. If the farmer wants to continue working with the watchman after the six weeks, then you can arrange a separate contract with the watchman once this project is finished. This project currently takes place during the upcoming short rainy season, with the planting beginning during August.

To learn about Security and Agriculture, it is important to compare people who have improved security to those who are farming under normal conditions. For this reason, once we have recruited a sample of ten farmers from each village in the sample who are interested in participating, we will enter all that data into a computer and randomly select the twenty villages where farmers will be matched with watchmen. We are not personally selecting the villages to be matched with watchmen and all villages have an equal chance of being selected.

This project provides only an improvement to security, not inputs or marketing services or anything else. There will be one watchman per farmer in selected villages. The watchman will protect the crops you choose to grow, all produce and profits are yours to keep. The watchmen will only be doing security work, not providing any other farm labour.

After we take your questions, we will break into groups by village and provide the names of ten farmers from each village who may be interested in this project, who will then be invited to a session to sign forms and answer a questionnaire.

Thank you for your time! I look forward to working together!