AMERICAN INSTITUTES FOR RESEARCH®

Evaluation of Plantwise–Kenya

Final Report

SEPTEMBER 2018

MAKING RESEARCH **RELEVANT**

Evaluation of Plantwise–Kenya

Final Report

SEPTEMBER 2018

AMERICAN INSTITUTES FOR RESEARCH®

1000 Thomas Jefferson Street NW
Washington, DC 20007-3835
202.403.5000

www.air.org

3869_08/18

Acknowledgments

AIR would like to acknowledge the contributions of the many organizations and people that contributed to the design, implementation, and production of this evaluation. First, AIR thanks DFID for providing the funding to carry out the evaluation and CABI for providing ongoing support for the design and implementation of the study. Thank you to our partners Dalberg and American University for data collection, design, and analysis contributions. AIR particularly acknowledges the participation of Ministry of Agriculture Livestock, Fisheries and Irrigation (MoALF&I, formerly MoA), Kenya Agricultural and Livestock Research Organisation (KALRO), University of Nairobi-Kabete Campus, and Kenya Plant Health Inspectorate Services (KEPHIS) staff in developing and scoring the test used to assess plant doctor knowledge. AIR would also like to thank the following groups for participating in interviews and providing valuable input and information to the study: Plantwise Steering Committee members; other members of participating agencies in the agriculture sector such as KALRO and KEPHIS; and, importantly, national and county MoALF&I stakeholders and partners who supported the study and all those who continue to support implementation of the program in Kenya. Finally, AIR would like to thank the thousands of farmers throughout the study areas in Kenya who participated in the research study over 4 years.

Contributors

The evaluation of Plantwise–Kenya is being conducted by American Institutes for Research (AIR) for CABI. Juan Bonilla (AIR) is the principal investigator for the overall evaluation. The team leaders of this report are Juan Bonilla (AIR), Andrea Coombes (AIR), and Nisha Rai (AIR), but many others made important contributions, and are listed below in alphabetical order:

AIR: Adria Molotsky, Paul Sirma, and Olatunde Olatunji

The suggested citation for this report is as follows:

American Institutes for Research. (2018). *Evaluation of Plantwise–Kenya: 36-Month Follow-Up Report.* Washington, DC: Author.

Contact information:

Juan Bonilla, jbonilla@air.org

List of Acronyms

AEA	Agricultural extension agent
AEO	Agricultural extension officer
AIR	American Institutes for Research
ANCOVA	Analysis of covariance
AWM	Area-wide management
CDA	County director of agriculture
CPO	Crop protection officer
DD	Difference-in-differences
FEO	Field extension officer
FGD	Focus group discussion
FHH	Female-headed household
ICT	Information and communication technology
IV	Instrumental variable
IPM	Integrated pest management
KALRO	Kenya Agricultural and Livestock Research Organisation
KEPHIS	Kenya Plant Health Inspectorate Service
KII	Key informant interview
LATE	Local average treatment effect
MLND	Maize lethal necrosis disease
MHH	Male-headed household
MOA	Ministry of Agriculture
NSC	National Steering Committee
PD	Plant doctor
PW	Plantwise
PW-K	Plantwise–Kenya
POMS	Plantwise Online Management System
RCT	Randomized, controlled trial
RSA	Research Solutions Africa
SAAO	Senior Associate Agricultural Officer
SCAO	Subcounty agriculture officer
SMS	Short message service
UoN	University of Nairobi

Contents

Tables

Figures

1. Introduction, Background, Objectives

This document presents the endline data collection and analysis for the evaluation of Plantwise–Kenya (PW-K). Plantwise (PW) is a global initiative that provides smallholder farmers with information on crop health. PW began working in Kenya in 2010 to gather, organize, and manage plant health information and disseminate that information to smallholder farmers. PW-K is implemented through three interconnected activities. First, farmers can access trained plant doctors through a network of locally run plant clinics. Plant clinics provide smallholder farmers with low-cost access to plant health information and diagnosis of plant health issues. Second, PW-K engages key stakeholders, such as the Ministry of Agriculture (MoA), extension programs, researchers, input suppliers, and regulators, in fortifying plant health systems through encouraging collaboration and coordination of activities. Third, a "Knowledge Bank" database serves as a repository for plant health information that assists with monitoring pests and diseases, promotes strategies for climate-change adaptation, and facilitates international trade.

American Institutes for Research (AIR), with support from American University and Research Solutions Africa (RSA), conducted a mixed-methods evaluation of PW-K. Qualitative methods assessed the changes that PW-K brings to the Kenyan plant health system and evaluated the fidelity of the PW implementation in the field. Quantitative methods identified farm-level impacts using an experimental approach based on the expansion of plant clinics in 2014 and 2015 in 13 counties in the country. The data collection plan involved baseline data and two rounds of follow-up data (at 12 and 36 months). Baseline and midline 12-month follow-up data were collected in 2014 and 2015, respectively, and have been analyzed in previous reports.[1] The final round of follow-up data was collected in 2017 and is analyzed in this report. The evaluation is based on the following four questions, identified at the onset of the study:

1. Does PW-K lead to stronger institutions for managing the plant health system—do these institutions expand knowledge availability, improve identification of new diseases and pests, improve response to pest and disease outbreaks (both at the farm level and nationally), and maintain numerous plant clinics? If so, can these changes be sustained over the long term?
2. How does the process by which PW-K is implemented influence the effectiveness of the program? Given this process, how do contextual factors affect program success?

[1] The baseline report, based on data collected in 2014, and the midline report, based on 12-month data collected in 2015, provide detailed description of the plant clinic context related to (a) a description of the agricultural situation in the study regions, (b) the major challenges to agriculture, (c) the way farmers obtain knowledge to address pest and disease issues, and (d) the differences in household and agricultural characteristics by household head gender. At midline, we found little change in the plant clinic context between baseline and the 12-month follow-up.

3. Does PW-K improve the wellbeing of farmers through improved pest and disease management, increased productivity, improved efficiency, and a rise in gross margins? Does this lead to increased income from agriculture and ultimately improved food security?
4. Are the additional costs of PW-K justified given the benefits the program provides?

This report discusses whether the endline data collection was successful in terms of gathering adequate data to address the research questions, and describes the data collected. The data was designed to provide final insights into the study's research questions based on three years of implementation.

The document is organized as follows: Section 2 provides the theory of change, Section 3 presents the overview of the study design, Section 4 answers the research questions using both qualitative and quantitative data collected as part of the evaluation, and Section 5 summarizes the key conclusions and discusses the implications of these conclusions for PW-K.[2]

[2] Both sections 2 and 3 were previously reported in the baseline and 12-month report and are included here for completeness.

2. Conceptual Framework: Theory of Change

This section presents the conceptual foundation for PW-K, previously introduced in the baseline and midline reports but revised to include key outputs we added to the theory. The theory of change presents PW-K's causal logic—how the activities of the program are intended to bring about improved development outcomes.

The theory of change starts by considering the **initial conditions**, which highlight the issues that PW-K aims to address. Smallholder farmers rely on crop production for income and food security; however, their ability to generate income and maintain food security is threatened by pests and diseases that reduce yields, as well as a lack of knowledge and access to information to address pests and diseases. Farmers require various options for addressing pest and disease outbreaks that consider context-specific agricultural conditions and socioeconomic circumstances (Danielsen et al., 2013). However, maintaining and updating information on the types of pest and diseases and how to address them is difficult in an agricultural landscape that is changing quickly due to globalization, market pressures, and climate change. Given this context, PW-K seeks to establish a better way to manage crop protection, to ultimately improve food security, alleviate poverty, and improve the lives of Kenyan farmers.

PW-K has three categories of activities: (1) institutional strengthening, (2) the Knowledge Bank, and (3) field activities. First, institutional activities include methods to provide support to and increase collaboration among participants in the national plant health system. The system exists through the MoA and the crop protection services it provides (e.g., research, agricultural extension services, regulation, data management, web services, a call center for farmers) as well as through affiliated stakeholders (e.g., universities, input suppliers, farm organizations). PW-K invests in establishing a support network for these local entities, seeking to facilitate information flow among them and coordinate action. Improved coordination is intended to increase the availability of information to farmers. In addition, it helps farmers benefit from coordinated responses and targeted messages that are immediate and delivered at scale.

The PW Knowledge Bank, the second component, is a free and open-access online database of locally relevant plant health information. It provides support to plant doctors who operate plant clinics, to extension workers, and to researchers. The bank provides diagnostic assistance, treatment support, and pest distribution data gathered from plant clinics, researchers, and international partners around the world. Investments from PW-K include establishing and hosting the database, along with providing ongoing advice to extension systems and national bodies. The online database is designed to increase the accountability and responsiveness of local organizations to farmers and guarantee the quality of information that farmers receive.

The final component of PW-K is to establish and maintain a network of plant clinics, which act as a physical interface between farmers and crop protection experts. The clinics are staffed by extension agents who work for the MoA, but receive additional training from PW-K to be plant doctors. Once trained, the plant doctors have the capacity to diagnose pests and diseases and to offer recommendations related to, for example, cultural practices (methods that do not use chemicals) or the use of chemicals. In general, the plant doctors seek to promote the idea of integrated pest management (IPM), which uses multiple approaches to pest and disease management that minimize hazard to people and the environment. While not excluding the use of chemicals, the plant doctors seek to ensure the rational use of these pesticides. Plant doctors are an alternative to suppliers who sell chemicals, in that they provide an expert opinion without any incentive to sell products. Plant clinics staffed by two plant doctors are set up throughout the country (59 clinics as of the beginning of 2014 and 90 additional clinics that started after 2014), often at agricultural markets, although occasionally at other locations where farmers congregate. The choice of location is generally based on making the clinics as accessible to the largest number of farmers possible. The technical support aspect of the PW approach is intended to strengthen farmers' knowledge of crops and crop health, as well as to open possibilities for agricultural diversification and production, as the costs of adopting new crops are reduced thanks to the presence of the plant clinics. The clinics are ongoing, and they seek to increase farmers' knowledge over the long term and help them to shift their behavior toward better crop protection practices.

The fact that the clinics represent the physical interface between farmers and crop protection experts highlights the broader role of plant clinics beyond the immediate intervention. First, the plant doctors are MoA staff, and they interact with not just PW clinic farmers, but other local farmers, other extension agents, and supervisors. Through these interactions they transmit knowledge that reflects their training as plant doctors, but also what they learn in clinics from farmers. This information exchange could also potentially influence the activities of local field agricultural extension offices. For example, the agricultural extension offices often hold field days for farmers to promote certain agricultural practices. The focus during field days can change based on the information obtained by plant doctors. Second, the plant clinics are designed to systematically collect information on the crops and associated pests and diseases reported by farmers. These data are entered in the Plantwise Online Management System (POMS), where the information is verified and ultimately used for analysis and to provide information (via the Knowledge Bank). The verification process involves carefully considering the data from clinics, which can then be used to provide additional training to plant doctors if issues are identified with their diagnoses and recommendations. The data also provide information for additional investigation if new diseases or pests appear to be emerging. The database itself serves as a key source of information regarding the prevalence of pests and diseases and emerging vulnerabilities in plant health. Overall, plant clinics play dual roles—addressing farmers' crop protection needs and stimulating broader institutional change.

Assuming all these activities happen and achieve the intended **outputs**—i.e., following through on the activities as intended—these three PW-K components are intended to lead to a set of **intermediate outcomes** (shown in Figure 2.1). These can be broadly categorized into (1) changes in the overall system for managing Kenya's plant health and (2) changes that result from farmers altering their behavior because of the plant clinics and general system changes.

The training of plant doctors and other members of the MoA should increase knowledge about crop protection throughout the entire crop protection system. The expectation is that advisory systems and regulatory systems for monitoring pests and diseases should improve with the shifts in management, expanded collaborative networks, and improved information gathering through the plant clinics and other information sources in the Knowledge Bank. Improvements in the overall management of crop protection services and the plant clinics should lead to greater quality in the provision of crop protection services if stakeholder behavior changes as a result of these investments.

With the improved overall service and availability of plant clinics, a behavioral response is expected from farmers as well. Farmers should respond by attending plant clinics and other related extension activities. The program should induce farmers to adopt new production practices if this occurs as planned and farmers internalize the information they obtain. Proper diagnosis should then lead to a better use of productive inputs. Improved input management should decrease crop losses and improve plant health and quality. It can also allow farmers to adopt new crops if concerns about pest and disease are limiting entry into the production of crops. Overall, improved plant health can increase sales and prices.

The behavioral responses by stakeholders and farmers rest on several assumptions, some of which are noted in Figure 2.1. The theory of change assumes that both plant doctors and farmers have a sufficient base of knowledge on which to build. It also assumes that stakeholders have internet access and can obtain the information in the Knowledge Bank. The theory assumes that farmers can and will travel to clinics, and that they will be able to use the information—in other words, that markets exist to obtain required inputs and that farmers have the resources to obtain inputs. If farmers find the information sufficiently beneficial to explore new production opportunities, the theory also assumes those opportunities exist and that the market can absorb additional products. These assumptions also demonstrate potential reasons why PW-K may not achieve its intended effects.

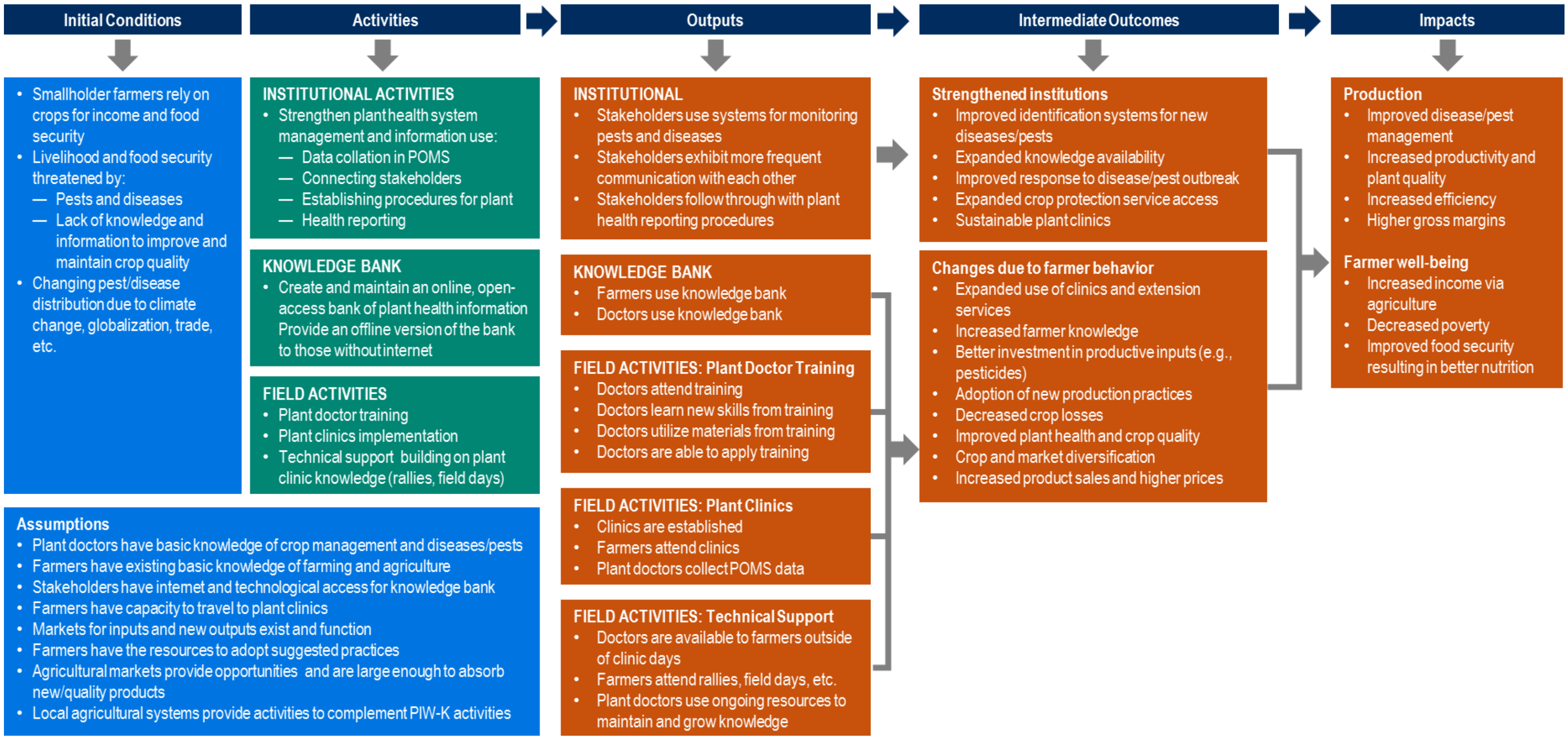
Taken together, stakeholders' behavioral changes would theoretically lead to stronger institutions for managing crop protection, higher quantity and quality of production, and greater income and food security for farmers. The stronger institutions would be seen in the context of a broader and continuously updated knowledge base, improved systems for identifying pest and disease outbreaks, and improved response to those outbreaks. The system would also have strong and sustainable plant clinics attended by well-trained and informed plant doctors. The program would then have impacts on agricultural production resulting from improved overall pest and disease management strategies, gains in efficiency, and higher gross margins. The program would also have impacts on farmer wellbeing, evidenced by increased income and decreased poverty, and improved food security and better nutrition. Figure 2.1 provides a graphical representation of the theory of change for PW-K.

The causal logic of the PW-K theory of change helped identify the evaluation questions. Two questions emerge directly: the impact of the program on the overall plant health system (Research Question 1) and the impact at the farm level (Research Question 3).

The impact of a program like PW-K is ultimately a function of how it is implemented within the plant health system. Understanding implementation is therefore a critical aspect of the evaluation, and investigated through research question 2. A key aspect of implementation in this case is the work of the plant clinics. Because they are the primary interface between PW-K and the farming community, the process by which they operate is critical to the success of the program. The process also influences the type of farmer the program reaches, and whether certain segments are excluded (e.g., women, poorer farmers, certain locations). The plant clinic process can also lead to variation in the effectiveness of the program through contextual factors such as the population density of a region or the type of farming present.

Another important consideration is the cost of PW-K. The program relies on existing resources in the Kenyan government; staff used in the program are already funded by government. The labor costs of PW-K are the opportunity cost of labor—the activities that government officers no longer do because of the program. Other programmatic costs relate to building capacity, implementing the plant clinics, and developing and managing the database, among others. Research Question 4 considers these costs in the context of the program's benefits.

Figure 2.1 Plantwise Theory of Change



3. Study Design

To answer the research questions, we used a mixed-methods approach that includes qualitative and quantitative methods from multiple sources. On the quantitative side, we analyzed both primary data and PW-K administrative data. The primary data collection included both a farm-level survey and a knowledge assessment of plant doctors. On the qualitative side, we collected information at the national and local (county) level and used both key informant interviews and focus groups. We also used cost data shared by CABI to conduct the cost-benefit analysis.

The logic of the data collection was based on PW-K's 2014 field presence and its plans for expansion. At the initiation of this study in 2014, PW-K operated in 13 Kenyan counties and had 59 plant clinics in operation. PW-K intended to expand the number of clinics within the country, with the expansion focusing on those 13 counties. The field work focused on the 13 counties covered by the program and exploited the intended expansion via a randomized phase-in approach to assess impact. In this section, we provide a summary of this approach. We start by describing the methods used for collecting and analyzing the data.

A. Randomized, Controlled Trial

The randomized, controlled trial (RCT) evaluation design we implemented allowed us to address two potential challenges that arise when estimating the effects of a free demand-driven program like PW-K. First, comparing the observed outcomes of farmers who attend plant clinics to those who do not may result in biased estimates of program impacts due to the self-selection of farmers into clinics. For example, if farmers who attend plant clinics are more motivated and ambitious than non-users on average, any improved production outcomes may be due to the farmer characteristics as well as clinic attendance. Alternatively, farmers attending plant clinics could be less experienced, have lower levels of agricultural knowledge, or be experiencing more pests and diseases relative to those who do not use the clinics. By randomizing the sites that received the program, we ensured that the farmers who live in the intervention areas are very similar on average to the farmers who live outside those areas; this approach addresses selection bias and allows us to estimate the true program impacts.

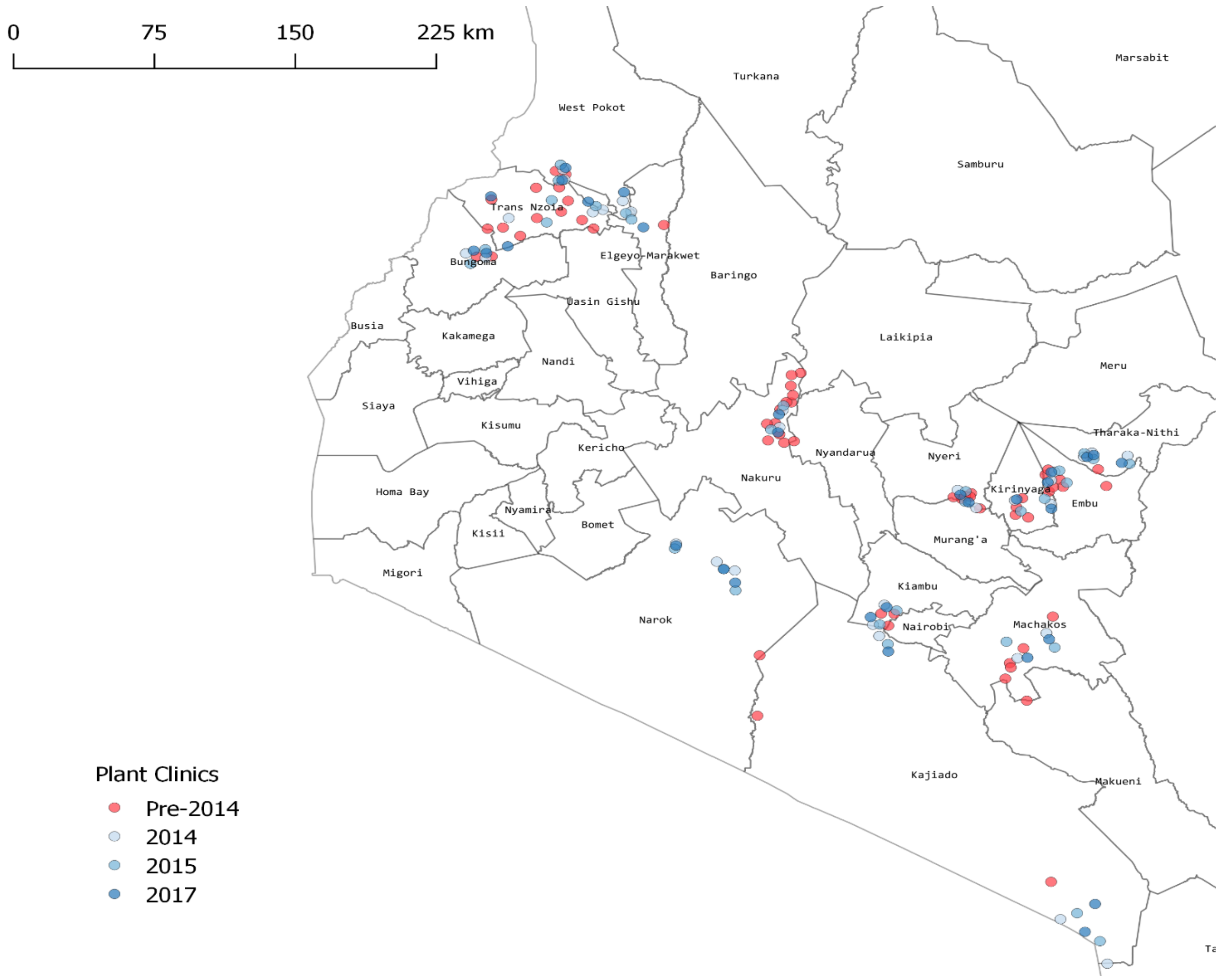
Second, PW activities may affect both plant clinic users and non-users in a variety of ways, such as through Plant Doctor knowledge sharing. Thus, any estimate of the effectiveness of the program needs to consider the fact that while some farmers in the community will benefit directly from the intervention, some will benefit indirectly through their interactions with direct beneficiaries and through an overall increase in plant-health knowledge in their community. Therefore, a fundamental characteristic of the evaluation design is that the program needs to be randomized at the community (cluster) level—and not at the individual level. With a cluster RCT, one must compare the outcomes of all those initially allocated to the treatment group to all those initially randomized to the control group, whatever their actual status in terms of clinic attendance, in order to capture the full impacts of PW on a given community.

To create the sample for the cluster RCT, county representatives from the 13 counties where PW-K currently operates were asked to identify 30 sets of three potential plant clinic sites with similar characteristics, for a total of 90 potential sites.[3] For each set of potential clinics, one site was randomly selected for inclusion in the program starting in August 2014, thus creating the initial 30 treatment sites and 60 comparison sites.[4] Selecting one of the three identified clinics, rather than a simple random selection of 30 clinics out of 90, ensures a similar distribution of treatment and control areas, which would not necessarily emerge from random selection. One additional site from each set of three was randomly selected for inclusion in the program after the first follow-up round of data collection in August 2015, and the final 30 sites received the program after the second (and final) follow-up round of data collection in August 2017. In other words, there were 30 treatment and 60 control sites from August 2014 to August 2015, there were 60 treatment and 30 control sites from August 2015 to August 2017, and all 90 sites were treated by August 2017.

Figure 3.1 presents the geographical location of all the plant clinics in Kenya at the end of 2017. The pre-2014 dots represent the 59 clinics that existed in Kenya before the current evaluation started. Those clinics are presented in the map for descriptive purposes but are not used in the evaluation. The blue dots represent the year in which the plant clinics used in this evaluation started, as discussed above.

[3] The average distance between an actual plant clinic and the closest control site is 5 km. Twenty-five percent of the clinics were located less than 2.1 km away from a control site. We discuss later how these distances between treatment and control sites may have affected the results.
[4] The random selection of clinics took place using a transparent process in the presence of county representatives, PW-K and CABI staff, and the evaluation team on April 30, 2014.

Figure 3.1 Location of Plantwise Clinics in Kenya in 2014



This delayed entry takes advantage of the natural expansion of the program to construct a comparison group for those who are selected to be treated over time. The randomization of plant clinic sites, as opposed to individual farmers, helps overcome concerns about not capturing the true program impacts, as the effects that PW has on farmers in treatment areas go beyond plant clinic attendance.

We implemented an “encouragement” design to increase the likelihood that farmers in treatment areas attended clinic sessions and to ensure a sufficient sample size in the treatment group. Surveyed farmers in treated areas received text messages with clinic hours and location, as well as additional information to encourage their use of the clinics. PW-K clinic coordinators from the treatment areas sent the messages in the days prior to implementing each clinic. In addition, CABI conducted sensitization activities in plant clinic areas in 2016 to increase farmer awareness of clinics.

To create a sample frame within the selected areas, potential program participants were identified through a census of farmers living in close proximity (within an average radius of 1.5 km) to the 90 treatment and control sites. Approximately 5,000 farm households were targeted for interviews as part of the census, which resulted in approximately 56 households from each of the 90 designated areas. Within households, we surveyed adult household respondents who were mainly responsible for the farming activities in the given household. PW-K, CABI, and the evaluation team identified the following criteria for inclusion in the census:

(i) Having between 0.25 and 10 acres of land for crop production (the size PW-K expects to benefit from the program)
(ii) Willingness to work with agricultural extension services
(iii) Willingness to provide mobile phone numbers

All identified households that did not meet the eligibility criteria were substituted with the immediate next household. Out of about 5,000 households originally targeted, approximately 75 did not meet the inclusion criteria, resulting in a final count of 4,925 eligible households.

Finally, we conducted a power analysis to determine a final sample size for the study that would allow the detection of meaningful program effects. This analysis indicated a need for around 31 household farms for each of the 90 sites, after accounting for up to 10% attrition over time. We randomly selected a sample of approximately 2,800 households from the census, equally distributed across all sites.

Identification Strategy for Impact Estimates

We use a cluster RCT to answer the counterfactual question of what would have happened to program beneficiaries in the absence of the intervention. With an RCT, the evaluation of the program's impact can be examined by comparing the mean of the treatment group with the mean of the control group. Such comparison works because when treatment and control individuals are randomly selected, the two groups are statistically equivalent, except for the fact that one group receives PW-K. Thus, all else being equal, the mean difference between the two groups represents the impact of PW-K. However, the extent to which comparing the means of the treatment and control groups is effective depends on whether the characteristics of the groups are perfectly balanced at baseline. In the baseline report (AIR, 2015) we showed that the randomization worked well in terms of creating equivalent groups—for most outcomes and control variables, the average characteristics of treatment and control groups were statistically equivalent. We tested all of the outcome measures and control variables for statistical differences between the two groups, using t-tests of differences in means across groups. Only four of the more than 140 variables analyzed were statistically significantly different between treatment and control at the 5% confidence level, as expected by chance.

In large-scale social experiments, it is typical to estimate program effects by using the experimental data within a longitudinal design. Using the values of an outcome of interest at baseline when estimating program impacts increases the precision of the estimator (i.e., reduces the standard errors of estimated impacts). A common longitudinal method for estimating program effects is an analysis of covariance (ANCOVA) estimation, which involves controlling for the baseline value of the outcome variable.[5] Specifically, we estimate the following ANCOVA specification:

$$Y_{ij,endline} = \alpha_1 + \beta_1 Treat_i + \beta_2 Y_{i,baseline} + \gamma_j + \varepsilon_{ij36}$$

Where $Y_{ij,endline}$ and $Y_{ij,baseline}$ are the outcomes of interest (e.g., any of the intermediate or final outcomes presented in Table 3.1) for farm i at endline and baseline respectively. $Treat_i$ is a dummy variable that takes the value of 1 if farm i is part of the treatment group in 2017 and 0 otherwise. Note that this implies that the estimated impacts in 2017 are the same irrespective of the year in which treatment occurred. We also control for γ_j, which are dummies for the 30 groups of clinics used for randomization (i.e., lottery fixed effects). No other covariates are included in estimation, as most of household-level characteristics collected at baseline (e.g., farmer's age, education, experience, household size) do not provide additional explanatory power once we control for the lottery fixed effects given the proximity of the clinic sites within a set of 3 clinics. Finally, standard errors of program impacts were clustered at the plant clinic level to account for correlation of farmers' outcomes within clinic areas and to handle potential heteroscedasticity of the outcomes.

Whenever possible, we report ANCOVA estimates for all relevant outcomes to maximize efficiency. Nevertheless, for outcomes for which we do not have the baseline value of the outcome for a large proportion of the farms in our sample, we report the difference between the treatment and control groups using data form the 36-month wave (i.e., single difference).

[5] ANCOVA exhibits large improvements in statistical power compared to using the more common difference-in-difference specification. The improvement in power is greatest when the autocorrelation is low –that is, when the baseline data have little predictive power for future outcomes, as is usually the case with agricultural outcomes (McKenzie, 2012).

Note that the single difference specification is the same as the one for ANCOVA except for the fact that it does not control for the baseline value of the outcome:

$$Y_{ij,endline} = \alpha_2 + \beta_2 Treat_i + \gamma_j + \varepsilon_{ij36}$$

Both the single difference and ANCOVA estimates are intention-to-treat (ITT) estimates—the average impact of being assigned to the treatment group (i.e., the average impact of being assigned to PW-K area) regardless of plant clinic attendance. Because the assignment was random and there is no evidence of differential attrition between the two condition groups,[6] we know that farmers assigned to PW-K are, on average, similar to farmers assigned to the control. This similarity ensures that our estimate of the impact is unbiased.

It is tempting to try to estimate the effects of PW-K for the subpopulation of farmers in the treatment areas who attended plant clinics by using the random assignment to a treatment area as an instrumental variable for plant clinic attendance—an estimate known as the Local Average Treatment Effect (LATE).[7] Institutively, the LATE measures the causal effect of attending a plant clinic for a very specific group, namely, the farmers who attended clinic sessions *only* because they were randomly offered access to the program by placing a clinic in their village.[8] The LATE estimate does *not* capture the impact of the program on other type of farmers, including those who would have attended a clinic under all circumstances, not just because a clinic was placed in their community (*always takers*).

Nevertheless, obtaining a consistent (i.e., unbiased) LATE estimate requires some assumptions that may not be satisfied in the case of PW-K (Duflo, Glennerster, & Kremer, 2008). Specifically, one ought to be concerned about violating the *exclusion restriction* assumption, a key assumption for consistently estimating the LATE of a program.[9] The exclusion restriction

[6] We discuss the attrition results in the next section.
[7] In the evaluation literature, an instrument is a variable that allows estimation of the causal effect of participating in a program activity (e.g., attending a plant clinic session) while addressing the fact that those who participate self-select into the program. A good instrument needs to have two characteristics. First, the instrument needs to be a strong predictor of the probability of program participation. In our case, this means that those assigned to a treatment area are substantially more likely to attend a plant clinic than farmers assigned to the treatment group. As shown in section Appendix 4, this condition holds in the evaluation of PW-K. Second, a good instrument satisfies the exclusion restriction assumption, which states that the only way through which the instrument (being in a treatment area) affects the outcomes of interest (e.g., yields) is through attending the program activity (attending a plant clinic). As discussed later, this condition may not hold in the evaluation of PW-K.
[8] This group of people is known as *compliers* in the evaluation literature. The other potential groups present in an evaluation are the *always takers*, those who always attend the plant clinics regardless of whether they are assigned to the treatment or control groups; the *never takers*, those who never attend the plant clinics regardless of whether they are assigned to the treatment or control groups; and the *defiers*, those who do the opposite of what they are expected to do according to their treatment status (attend plant clinics if they are assigned to the control group, and not attend the plant clinics if they are assigned to the treatment group).
[9] In Appendix 4, we describe the LATE method in more detail.

assumption implies that the instrument (i.e., being in a treatment area) does not affect the outcomes of those farmers who, despite living in treatment areas, decide not to attend plant clinics. However, one fundamental characteristic of PW is that trained plant doctors interact with farmers, other extension agents, and supervisors not only thought plant clinics, but also outside the plant clinics. Moreover, farmers in treatment areas who do not attend plant clinics can also benefit from the fact that their crops may be healthier if neighbors who attend plant clinics have healthier crops as a result of clinic attendance. Thus, the fact that PW activities may directly affect not only plant clinic users, but more importantly non-users of plant clinics, in treatment areas implies that the estimation of a LATE can be severely biased. While the violation of the exclusion restriction assumption cannot be tested empirically, its implications need to be carefully considered to avoid reintroducing biases in the estimation of program impacts.

In sum, to assess the impact of PW-K, the cluster RCT design we use takes into account key features of the program, including the fact that PW is a demand-driven intervention, with public-good characteristics (any farmer can attend a plant clinic at no cost regardless of where they live), and where program benefits in a community are realized even beyond plant clinic attendance. The evaluation design allows us to estimate the ITT of the program. In the context of PW-K, the ITT is the actual parameter of interest, as we are interested in the effectiveness of an agricultural extension model that intends to serve entire communities. Thus, any estimate of the effectiveness of the program needs to consider the fact that while some farmers in the community will benefit directly from the intervention, some will benefit indirectly through their interactions with direct program beneficiaries and through an overall increase in plant-health knowledge and status in their community. Therefore, our discussion of program results in section 4C focuses on the ITT estimates of PW. In Appendix 1, we present the LATE results for the different outcomes of interest, which need to be interpreted with caution for the reasons cited above.

B. Quantitative Data Collection

The Farmer Questionnaire

We designed farm-level instruments to collect data on intermediate outcomes related to short-term, welfare-improving behavioral changes by farmers:

1. Investing in better production inputs
2. Adopting new practices

3. Improving disease/pest management (such as better use of pesticides)

To measure the effects of the program on final outcomes, we collected detailed information on crops that were cultivated in an area larger than 125 square meters (or 1/32 acre). These final outcomes include crop production amounts and market values, which—along with input expenditures such as fertilizers, pesticides, and labor—enabled us to estimate program effects on production value, yields (production per unit of area), and gross margins.

We also collected information about farmers' awareness, usage, and support from PW-K and other extension services to investigate the fidelity of program implementation in the field.

We collected information for 2,828 farmers in 2014, for 2,504 in 2015, and for 2,563 in 2017. We discuss the attrition rate and its implications below.

Table 3.1 presents the most relevant intermediate and final outcomes we collected in the survey. The table describes the type of variable—whether it is a categorical or continuous variable and the units of measurement— as well as the level at which the variable can be constructed. We analyzed outcome variables at three different levels: the crop level (C); annuals (A) or perennials (P) (after aggregating crop data by plant life-cycle type);[10] and when all crops are aggregated at the household level (H). These levels help better characterize the agricultural conditions of households with multiple crops. Note, for instance, that the value of a specific variable (say, gross margins) only varies by level when the household has more than one crop. Having only one crop yields the same value at all levels, since the crop level is the same as the household level.

Table 3.1 Description of Selected Outcome Variables

Outcome	Variable type	Variable level
Intermediate outcomes		
Cultural practices		
Crop rotation (and reasons why), early planting, intercropping (and reasons why), removal of plant residue, use of improved planting material, planting resistant varieties, use of certified planting material, crop monitoring (including times	Yes = 1, No = 0	C, A

[10] Some common annual crops produced by farmers in our sample are maize, beans, potatoes, kale, and tomatoes. The most common perennial crops are bananas, avocados, coffee, napier, woodlot, and mangoes. We present the results by county in section 4C.

Outcome	Variable type	Variable level
inspecting crops for infestation), use of trap cops, burning crop residue or trash to control pests		
Planting resistant varieties, use of improved planting material, intercropping (and reasons why), weeding, removal of infested/damaged material, crop monitoring (including times inspecting crops for infestation), removing plant residue, use of trap cops	Yes = 1, No = 0	C, P
Inputs		
Value of seed planted (imputed)	KHS and KHS/acre	C; A, P; H
Organic fertilizer used	Yes = 1, No = 0	C; A, P; H
Quantity of organic fertilizer used	Kg/acre	C; A, P; H
Inorganic fertilizer used	Yes = 1, No = 0	C; A, P; H
Quantity of inorganic fertilizer used	Kg/acre	C; A, P; H
Value of inorganic fertilizer used	KHS and KHS/acre	C; A, P; H
Pesticide used	Yes = 1, No = 0	C; A, P; H
Number of pesticide applications	Integer	C; A, P; H
Quantity of pesticides used	Kg/acre	C; A, P; H
Value of pesticides used	KHS and KHS/acre	C; A, P; H
Labor days used for pesticide application	Integer	C; A, P; H
Protective measures for pesticide application (use of protective items, disposal, storage)	Yes = 1, No = 0	H
Biological crop protection used (e.g., use of microbials, macrobials, traps, plant extracts)	Yes = 1, No = 0	C; A, P; H
Value of biocontrol used	KHS and KHS/acre	C; A, P; H
Total family labor days	Number of days	C; A, P; H
Total paid labor days	KHS and KHS/acre	C; A, P; H
Cost of paid labor	KHS and KHS/acre	C; A, P; H
Production diversity		
Number of crops produced	Count	A, P; H

Outcome	Variable type	Variable level
Final outcomes		
Yields	Weight units/acre	C; A, P; H
Total value of harvest	KHS and KHS/acre	C; A, P; H
Gross margins per unit of area	KHS (see formula)	C; A, P; H
FANTA household dietary diversity measure	Count	H

Note. C = Crop; A = Annuals; P = Perennials; H = Household level. KSH: Kenya Shillings.

As shown in Table 3.1, almost all outcome variables can be analyzed at the crop level (C), including cultural practices, inputs used (e.g., seeds, fertilizers, pesticides, labor), and production variables (e.g., value of harvest and gross margins). In addition, most outcomes can be aggregated into either annuals (A) or perennials (P), and then totaled by household (H). The most notable exceptions are the two indices on production diversity, which are more meaningful at the household level.

Intermediate Outcomes

The most relevant intermediate outcomes were indicator variables for cultural practices used, use and value of organic and inorganic fertilizers, use and value of pesticides used, use and value of biological protection, and quantity and value of labor used.

Final Outcomes

Beyond these immediate indicators, the survey included details on all the crops produced by farmers, regardless of land area, which enabled us to assess whether plant clinics have an impact on crop production diversity. Along similar lines, the survey included details about food consumption within households to assess whether production diversity or higher income led to more dietary diversity, which is a proxy for food security. We also collected information on crop production, which enabled us to estimate program effects on yields (i.e., quantities per unit of cultivated area) and impacts on gross margins (G) when combined with data on input expenditures on fertilizers, pesticides, and labor. Gross margins for crop *c* in household *h* is defined as:

$$G_{ch} = P_c * Q_{ch} - P_L * L_{ch} - P_b * B_{ch} - P_{QP} * QP_{ch} - P_F * F_{ch} - P_S * S_{ch}$$

where P_i is the price per unit *i*, including quantity harvested (*Q*), hired labor (*L*), units of biological control used (*B*), pesticide used (*QP*), fertilizer used (*F*), and seed planted (*S*).

Note that to calculate values such as the value of pesticides or harvest, we multiplied the price by the quantity used. All quantities were from the farmer survey. Following common practice, missing prices, on the other hand, were sometimes imputed using the median reported price for a given commodity or crop at the plant clinic area level. If too few observations were available for a specific crop and input within a plant clinic area, we imputed the median prices at the next available aggregation level—that is, at the county level, and if that level was unavailable, at the national level. In addition, we converted all area variables to acres to express all monetary values per acre in order to facilitate the comparison of farm households with different land areas.

Control Variables

We also collected a set of variables to characterize households in the sample. The variables we collected, including head of household characteristics, relate to the socioeconomic conditions of the household. These variables include the age, gender, and educational attainment of the household head; household demographic composition, such as number of household members and age ranges; housing conditions, such as the materials used for walls, floors, and roofs; and a description of the main assets owned by the household and household access to public services (e.g., water and electricity sources). Note, however, that, as discussed earlier, the regression models we used to estimate program impacts always control for the dummy variables for the 30 groups of clinics used for randomization (i.e., lottery fixed effects).[11] By controlling for the lottery fixed effects, we are able to capture most of the variation in the socioeconomic characteristics of farmers within a given set of three sites. As a result, once we control for the lottery fixed effects, the household level characteristics discussed above provide no explanatory power for the outcomes of interest (i.e., they are not statistically significant), and adding them hurts the statistical efficiency of the estimated models.

Attrition

Attrition within a sample occurs when households from the baseline sample are missing in the follow-up sample. Mobility—the dissolution of households, death, and divorce— can cause attrition and make it difficult to locate a household for a second data collection. Attrition causes problems in conducting an evaluation because it not only decreases the sample size (leading to

[11] Recall that randomization was done after selecting 30 sets of three potential plant clinic sites and then randomly assigning the order in which each one of the three sites per set would receive the program over time. In other words, program randomization was performed by running 30 individual lotteries.

a less precise estimate of program impact), but may also introduce selection bias to the sample, which will lead to incorrect program impact estimates or change the characteristics of the sample and affect its generalizability.[12] In particular, differential attrition occurs when the treatment and control samples differ in the types of individual who leave the sample. Differential attrition can create biased samples by eliminating the balance between the treatment and control groups that was achieved through randomization at baseline.

We investigate differential attrition at the 36-month follow-up by testing for similarities at baseline between treatment and control groups for all non-missing households. Approximately 9.8% of the sample was lost to attrition. The attrition was uniformly spread throughout all 90 study sites, so that on average, in each site three of 30 respondents were not interviewed. The overall response rates are balanced between treatment and control groups. Table 3.2 shows the household response rates by treatment status at the 36-month follow-up for each county.

Table 3.2 Household Response Rates, by County

County	Control	Treatment	Total
Bungoma	91.3%	90.3%	172
Elgeyo-Marakwet	95.1%	93.5%	263
Embu	94.6%	96.9%	185
Kajiado	87.8%	82.3%	245
Kiambu	75.2%	90.5%	151
Kirinyaga	89.7%	84.4%	167
Machakos	86.6%	93.7%	169
Nakuru	83.3%	87.3%	160
Narok	95.7%	94.6%	265
Nyeri	92.1%	87.3%	171
Tharaka Nithi	85.2%	93.5%	248
Trans Nzoia	94.4%	93.7%	177
West Pokot	94.5%	93.4%	177
Overall	89.8%	90.9%	2,550

Note. Total number of baseline observations is 2,828. Total number of 2017 endline observations is 2,550.

The reasons for attrition are similar to the reasons reported at midline, and include the following: (1) target respondents unknown in the target areas because of wrong names or other

[12] What Works Clearinghouse (http://ies.ed.gov/ncee/wwc/documentsum.aspx?sid=19)

information captured in 2014 (approximately 48% of attrition cases); (2) target respondents refused to participate (24% of cases); (3) target respondents relocated to other counties (19% of cases); (4) target respondents were traveling at the time of data collection (3% of cases); (6) target respondents were duplicated at baseline (2.6% of cases); and (7) target respondent passed on and household was deserted (1.5% of cases).

In our initial power calculations, we accounted for 10% attrition to determine the study sample size. Thus, the observed attrition rate does not compromise our ability to detect meaningful program impacts. Also, at baseline, we collected information for 2,828 households, even though the power calculations indicated we needed 2,800 households. Moreover, although we still had a large enough sample size to detect meaningful impacts, we tested whether the benefits of randomization were preserved at follow-up due to attrition.

We did not find any significant differential attrition at the 36-month follow-up, meaning that the households that dropped from the sample are not observationally different from those that remained in sample, which means we are still able to attribute any observed differences in outcomes between treatment and control households to the program. Appendix 2 reports the results for the tests on mean differences for all of the outcomes and controls. These tests examine any statistical differences at baseline between the treatment and control households that remain in the 36-month follow-up analysis. There are minimal differences between these two groups at the 36-month follow-up. Out of 132 variables tested, differences existed in only 17 of the variables. Furthermore, most differences (13 of the 17) were of weak statistical significance at the 10% significance level. Finally, the experienced attrition does not have major implications on statistical power to detect program impacts given that the evaluation uses a cluster-randomized design, which means that most of the power is driven by the number of clusters (i.e., plant clinic catchment areas), and not by the number of observations per cluster. Thus, statistical power is almost unaffected due to attrition, because we only lost a few observations per cluster as opposed to losing entire clusters.

The similarity of the characteristics of people missing in the follow-up sample between treatment statuses allays concern that attrition introduced selection bias. Thus, the study maintains strong internal validity created through randomization, enabling estimated impacts to be attributed to PW-K rather than to differences in the groups resulting from attrition. While there are differences between farmers that attend the clinic and those that do not, our main impact estimate is an intention-to-treat estimate that relies on the randomized assignment of the clinic areas, the benefits of which are preserved due to minimal attrition.

Plant Doctor Selection, Assessment, and Implementation

To investigate whether plant doctor training improves the knowledge of agricultural extension agents (AEAs), we developed a knowledge assessment in collaboration with PW-K, CABI, and the University of Nairobi (UoN). We conducted the assessment in 2014, 2015, and 2017. The assessment participants came from three different groups of extension officers: (1) all plant doctors who started working in the plant clinics before 2014 and were trained before the 2014 assessment; (2) agricultural agents selected to become new plant doctors over time for the new plant clinics that started after September 2014; and (3) a set of comparable AEAs, who have not received training from PW-K, but work in areas close to plant clinics and have similar observable characteristics to plant doctors. This last group serves as a control group to assess the effects of plant doctor training. By testing the different types of extension agents over time, we can estimate the causal relationship of the PW trainings on plant health knowledge.

The plant doctor assessment consisted of two parts: (1) a multiple-choice test with 50 questions, which all participants should have been able to complete within 75 minutes; and (2) a short-answer section, which participants should also have been able to complete within 75 minutes. Each section was worth 50 points, for a total of 100 points. The multiple-choice questions were related to the knowledge necessary for diagnosis and providing relevant recommendations; the questions were designed to be relevant in Kenya. Questions were easier at first and became incrementally harder. The short-answer questions were more comprehensive and were intended to simulate the conditions faced by plant doctors. Questions incorporated diagnosis, recommendations, and potential behavioral responses. Some short-answer questions had multiple parts, focusing first on diagnosis and then on recommendations. Again, questions had different difficulty levels to ensure scores captured variations in plant health knowledge. Finally, we asked PW-K, CABI, and UoN to design the 2015 and 2017 assessments based closely on the 2014 assessment to ensure the tests were horizontally equated[13] to facilitate the interpretation of score gains over time.

We provide methodological details and descriptive statistics of the sample used for analysis in Section 4 of this report.

[13] Test equating is the process of determining comparable scores on different forms of an exam. Horizontal equating allows comparison of the scores of two tests administered at two different points in time, where the tests are different to avoid practice effects.

Cost Analysis

To assess whether the benefits of PW-K justify its costs, we monetize estimated impacts of the program and compare them to the costs of implementing PW-K. Identifying costs associated with the project is not trivial, as costs include the resources provided by CABI and PW-K (including the costs of running the program by the government). To identify the costs, the ingredient method is used, in which every ingredient that could change an effect resulting from an intervention is considered. The information on the types of costs associated with the program was shared by CABI. After identifying these ingredients, we compare the actual costs from CABI and PW-K to the benefits (program impacts). We present the details of this exercise in Section 4D.

C. Qualitative Data Collection

Qualitative research aimed to provide in-depth analysis of and insights into the impact of PW-K. We conducted key informant interviews (KIIs) and focus group discussions (FGDs) at the national and county levels. Given that the benefits of KIIs and FGDs tend to diminish with each additional interview (as saturation occurs), we focused the field interviews in 4 of the 13 counties where PW-K operates. With the objective of capturing a broad range of experiences, CABI and the MoA selected four counties—Trans Nzoia, Nakuru, Machakos, and Kirinyaga—to represent the geographic and agroecological diversity (e.g., irrigation, crops, land size) of existing clinics. We gathered information at three points in time: baseline (all interviews), 12 months after baseline (county interviews and treatment FGDs), and 36 months after baseline (National Steering Committee [NSC] members and county-level stakeholders). This design gave insight into the plant health system with existing clinics, without clinics (before areas received the PW-K intervention), and then again after communities adjusted to having the clinics. It also allowed an assessment of the sustainability of any changes in the plant health system.

We conducted a stakeholder analysis in April 2014 to identify the appropriate parties to interview at the national level. An important part of the PW approach at the national level is the National Steering Committee, which gathers representatives from a variety of organizations in Kenya's plant health system to provide program oversight and input. The committee also ensures high-level stakeholder involvement, which ideally encourages program buy-in. The team interviewed all members of the NSC, as well as at least one other individual from the NSC member agencies to understand whether PW-K's influence extended beyond NSC members in those organizations. The non-NSC individuals were aware of PW-K, but are not active participants in the NSC. Finally, we also interviewed 10 representatives of organizations that are part of the Kenyan plant health system, but not directly involved in PW-K. In total, we conducted 34 national-level KIIs at baseline. We did not conduct national level interviews at midline; we interviewed 11 NSC members at endline.[14]

At the county level, we conducted 20 interviews: seven officers from the research institute Kenya Agricultural and Livestock Research Organisation (KALRO) and the plant health organization Kenya Plant Health Inspectorate Service (KEPHIS); and 10 county officers, including four desk officers, two subcounty agriculture officers, and four clinic cluster coordinators. We also interviewed three agrodealers who were available to discuss whether and how PW-K influenced their businesses. Cluster coordinators selected the agrodealers based on proximity to clinics, but the agrodealers were different at baseline and endline.

Finally, cluster coordinators helped identify farmers that had attended clinics and plant doctors to participate in 16 FGDs at baseline, midline, and endline. At endline, all farmers in FGDs had attended a clinic at least one time. One set of FGDs included male plant doctors from each of the four qualitative study counties, while a second set of focus groups included female plant doctors from each of the four qualitative study counties. Similarly, we interviewed one group of female farmers and one group of male farmers in each qualitative study county, for a total of 16 focus groups across the four counties. We separated men from women to understand differential experiences by gender for farmers and plant doctors.

The KIIs and FGDs used semistructured protocols that included specific questions about the program, but also allowed for adding probes in a more free-flowing conversation to capture information that might otherwise be missed. The questions were intended to allow the participants to steer the discussion, while still ensuring that they stayed on relevant topics. At baseline, the guides focused on gaining an understanding of whether stakeholders implemented the activities as expected in the theory of change, whereas the guides at mid- and endline followed up on key topics of interest, such as collection and collation of agricultural data, and whether men and women had different experiences with the clinics and how. We summarize the sampling for endline in Table 3.3.

[14] NSC members who were not interviewed either could not be reached or missed the interview after multiple attempts to schedule. Members of non-NSC organizations from baseline were contacted to ask if they had engaged with PW-K at all since baseline. None of the members responded that they had, so we did not interview this group.

Table 3.3 Endline Sampling for Qualitative Data Collection

Subject	Description	Key informant interviews	Focus group discussions	N
Farmers	1 group of 6–8 male farmers from each of the study counties; 1 group of 6–8 female farmers from each of the study counties		•	8
Plant doctors	1 group of male plant doctors from each of the study counties; 1 group of female plant doctors from each of the study counties		•	8
Agrodealers	Agrodealers from the selected counties	•		3
County-level KALRO and KEPHIS	Officers from KALRO research institute and KEPHIS plant health organization	•		7
County-level government officers	Subcounty agricultural officers (SCAOs) (2) Plant Clinic cluster coordinators[15] (4) County desk officers (4)	•		10
NSC members	Members of the National Steering Committee	•		11

The qualitative research team collected endline data from July through September 2017. Before analyzing data, the team deidentified participant information, clarified shorthand in interview notes, and reviewed transcripts for any inaudible comments and corrected the files.[16] The team developed broad categories that represented themes as they emerged during the data collection. These categories formed a coding scheme that aimed to separate raw data into large categories, and then further into more nuanced subcategories within each theme. The team defined each node to ensure consistency across coders and over time.

While coding qualitative data into NVivo Data Analysis Software, coders met to discuss new codes and revisions to the coding scheme. The analysis used multiple triangulation techniques (Denzin, 1978), including methodological triangulation (Lincoln & Guba, 1985; Guba & Lincoln, 2005), which helped promote the integrity of the overall research while also generating sufficient data for describing the interventions. The study team's objective was to ensure that all facets of the research (e.g., data collection, data management, data analysis, and reporting) were systematically coherent, to ensure the credibility of the findings. This approach to the qualitative analysis formed the basis of the insights provided in the following sections.

[15] Cluster coordinators are also trained as plant doctors.

[16] KIIs were conducted in English, while FGDs were conducted in Kiswahili. FGDs were transcribed and translated from Kiswahili to English, using a forward/backward translation process.

4. Results

Evidence from the impact assessment firmly establishes that PW-K had a number of positive effects on the Kenyan plant health system. First, PW-K appears to have been an impetus for institutional change, increased awareness of plant health issues, and altered the manner in which the government addresses crop protection. Second, the process through which PW-K is implemented is innovative and comprehensive. It improves knowledge at multiple levels through improved training for extension officers; accessible diagnosis for farmers; and data collection to help understand where diagnosis could be improved in the short term, and where the system should address problems in the long term. In terms of knowledge improvement, a plant doctor assessment found that trained extension agents scored significantly higher than non-trained extension agents. At the farm level, we find improvements in the use of cultural practices and inputs for farmers in areas with access to plant clinics. One key finding of the evaluation is that maize farmers in the treatment group experience a large and statistically significant increase in the value of maize per acre over farmers in the control group. Lastly, we assessed how the benefits compare to the costs of implementing the program. The estimated profitability measures show that PW-K provides a good value for money, in clear contrast with recent evaluations of similar extension programs, which have not exhibited impacts on farmer-level outcomes. We organize this section by the four primary research questions that guide the study.

A. Does PW-K lead to stronger institutions for managing the plant health system?

PW-K has facilitated institutional coordination in the plant health system and subsequently improved the likelihood of identifying pest and disease outbreaks and responding to them in a timely manner. Stakeholders perceived that PW-K has altered how farmers interact with MoA entities at the local level. Plant clinic activities have become a regular part of the MOA activities in the counties where PW-K operates; nevertheless, PW-K continues to be affected by challenges with the funding and organization of government systems after devolving administrative government responsibility from the national level government to the county level. The data indicate that cross-organization engagement does not happen regularly. In addition, PW-K is still not well known within some organizations, which causes a key gap in engaging with the public sector to harmonize its processes with the work and data that PW-K provides. This section provides qualitative endline results on the intermediate outcomes related to institutions, including improvements in (1) county-level advisory services, (2) data collection and reporting, and (3) pest and disease identification and response.

Improved County-Level Advisory Activities

PW-K is a well-regarded program among county- and national-level agriculture officers that provides concrete activities and is a primary focus of employees and users in the system. In addition to improving the knowledge of plant doctors as a result of their training (see plant doctor assessment results in section 4B), PW-K has also made it easier for farmers to access services through centrally located clinics which specifically address their challenges with pests and diseases. Despite the overall improvements in advisory services as a result of PW-K, county officers perceived that the MoA prioritizes crop protection only *after* identifying a pest or disease outbreak, meaning that PW-K lacks support in terms of financing logistics (e.g., airtime, internet access, transport), sufficient staff, and non-PW-K extension services.

Respondents said county offices of agriculture had limited capacity to support PW-K or regular advisory services because of funding and coordination challenges after devolution. County officers said they rarely received the funding from the national-level MoA to support the activities of PW-K under the devolved system of government. As a result, plant doctors said they did not receive county-level in-kind contributions for transport allowances and air time to complement PW-K. A respondent from Nakuru said, “The amount of funding allocated to extension services has been low over the period that the governments have been devolved.” Because they were underfunded, counties typically allocated little of their budget toward agriculture—where projects were not as highly visible as those in other sectors. Respondents said MoA finance for pest and disease management, specifically, came from an emergency fund as opposed to yearly funding, which put a noticeable strain on activities at the county level. Emergency funding was also processed too slowly to serve as a response to early warning or to respond to outbreaks once they had already spread.

National-level MoA priorities and employee workload affected whether clinics could run as intended. For example, a cluster coordinator said value addition was currently a priority, but did not think extension services were a priority “because we are getting fewer and fewer by the day, and there is no recruitment that has been done or the county is planning to do.” Another cluster coordinator said he did not think pest and disease management was a priority, while a county-level KALRO representative said, “I would say yes, that crop pests and diseases management is a priority to the county—though it is a silent priority.” County-level priorities are typically a reflection of broader and widely publicized national-level priorities.

Farmers expressed a lack of support from MoA officers outside of PW-K. One farmer said, “I wanted to say that we are unsure what happened to the government… they have not to come to the grounds.” MoA officers expected that farmers would demand advisory services that allow them to increase their productivity, but transaction costs for accessing government advisory services—such as distance to the office and the chance that an AEA is not present—hinder their opportunity to use them. Respondents thought food security was a priority for the government, so it may help PW activities and funding for crop protection activities to explicitly link crop protection to food security.

Improved Data Collection and Reporting

MoA officers said that each county reports on production indicators to the national MoA, but indicated that PW-K reporting systems filled a notable gap in data capture and reporting, particularly on pest and disease prevalence and diagnosis. County-level officers recognized the value of POMS data for tracking disease outbreaks, and higher-level external stakeholders, such as national-level officers, inquired about the data from the system on occasion. This section describes how PW-K data enhance the MoA-required data.

MoA Data Collection and Reporting

County officers explained that, before devolution, extension agents reported on indicators such as crop yields, losses, and production to the community, province, and then national level offices; however, the information was not systematically collected, and extension agents were not experts on the information they were collecting. Respondents indicated that under the devolved system, county level officers are still expected to report on yields, losses, and production to the national level, though they said information is still not systematically collected across counties. A county-level report provided by a national level MoA officer contained information on production for one county, but no information on pest or disease status or prevalence.

County officers were unclear on how much reporting on production to the MoA was actually taking place, and multiple stakeholders said the data was inaccurate. Descriptions of reporting varied by county, but most county-level officers broadly said they, “report *some* to the national level.” Although MoA officers at the national level said counties reported production yields and losses as described above, they confirmed said there were no official national-level requirements. Though some reporting may be required at some of the district and county level offices, it is evident that there is no system for utilizing the information in the reports to respond to ground-level problems related to pests and diseases or general challenges with funding and operations in agriculture.

PW-K Data Collection and Reporting

County-level officers consistently indicated that PW data reporting was timelier than government systems, and that PW-K data were more accurate than agricultural data collected from county-level MoA officers on production, yields, and losses because of the validation process. National- and district-level officers said the primary source of information on pests and diseases is from PW-K. For example, one desk officer said that he analyzes "not so much" data aside from the POMS:

> I only do analysis for the POMS because it is easier to work with. The reason I don't like analyzing the [production] information that we collect at the county level is because there is lack of uniformity on how the data is organized; some [counties] do not complete some parts and so it becomes difficult to analyze some of that data.

Multiple stakeholders said they valued PW-K data for tracking disease outbreaks, although the extent to which PW-K was responsible for identifying outbreaks was unclear. One sub-county agriculture officer (SCAO), among many others, said,

> We are able to tell from the POMS which part of the year a pest is more prevalent that other times, and if we do it over a period of time we are able to predict when a pest is likely to attack and prepare farmers with the control and management measures that they would require.

Awareness of the existence of POMS data also seems to have increased since midline; national-level officers inquire about the data that come from the system, although they are not versed in how to use the system.

County-level officers say they are using POMS and see the value of POMS data for tracking disease outbreaks; however, POMS data are still not being widely used. For example, one cluster coordinator said,

> The last time we used it we were doing a presentation to the county on the prevalence of certain pests and diseases to the governor and other stakeholders… I think that was

> done in 2015. We also presented issues on the gender of farmers who attend the plant clinics.

Districts lack a consistent system to employ the data systematically. Respondents attributed low use of the POMS data to challenges that had simple fixes, including not having passwords and not having a basic understanding of how to use the data.

Improved Identification of Pest and Disease Outbreaks

Plant clinics were mentioned most frequently as the primary way that stakeholders identify pests and diseases. PW-K data are a unique trove of local knowledge on pests and diseases that does not exist elsewhere in the country, and that contributes to the better known and widely used PW Knowledge Bank.[17] Other means of finding out about pests and diseases included through the media, the national government, and "telegram," a system used by plant doctors exclusively that seems to be relatively effective in communicating across counties. Plant doctors in five of the FGDs referenced their use of telegram as being a convenient and useful way of inquiring about a pest or disease to other plant doctors who may have encountered a pest. One female plant doctor also said, "If anything [new comes out] in research, we are informed through the forums we have like telegrams and tablets. That information [is] updating our knowledge in the field of agriculture extension." Another female plant doctor from a different county said there was nothing like the telegram medium for consulting on plant health problems, and that "so many of us" are utilizing it.

Most respondents said there is no established government system for detecting or reporting new pest invasions. When asked about whether there is a system in the county for detecting a new pest invasion, a cluster coordinator said, "The only system is the plant clinics, because we receive samples from the farmers and we get to learn about new pests in the plant clinics. KALRO also participates in alerting officers as well as farmers on a new pest invasion." A few officers mentioned surveillance activities, but it seemed these activities only occur on a quarterly basis at most. One desk officer described the process of surveillance, which relies primarily on plant clinic data, with some field visits:

> In case of an outbreak, we go together as a team, look at the severity, and discuss what could be the feasible options for solving the problem and how to manage and advise the county management. When there is no outbreak, we do

[17] As described in the conceptual framework, the Bank provides diagnostic assistance, treatment support, and pest distribution data gathered from plant clinics, researchers, and international partners around the world.

> quarterly visits to the field and focus on status with pest infestations. Ideally, we are supposed to do quarterly visits, but because of limitations of resources we don't normally do that. What we rely on mostly is the plant clinic data. We look at it, see the pest situation, and make necessary decisions based on the data.

In addition, KEPHIS personnel indicated that they checked one or two samples from the plant clinics over the past year, a process that seemed to take months if the pest or disease was not identified as a priority. A representative from KALRO suggested, "The challenge is that we lack enough surveillance system[s] that will help us detect new pests before farmers complain or lose to the pest. I think we need a more organized surveillance system to be able to forecast presence of a new pest." The representative did not speak to PW-K's specific role in surveillance.

Improved Systemic Response to Pest and Disease Outbreaks

Systemic response to pests and diseases seems to be more organized because of PW-K systems and the availability of POMS data. One county-level desk officer said,

> When we are able to analyze information from the field we are able to tell if there is an increase in a certain pest or disease in a certain area—so our officers are giving feedback. Currently we have a challenge with fall army worm infestation, and from the reports we get from our plant clinics, we are able to map out areas that are heavily infested by the pest and act very quickly. Our officers are also very fast in doing diagnosis and giving feedback to the farmers on what should be done, how it should be done, and when.

However, response to pests and disease outbreaks continues to be slow because of the challenges with coordination at the government level and systemic reporting. For example, although national-level stakeholders assumed extension agents and plant doctors went to KALRO for help with diagnoses, KALRO representatives said they rarely received samples from the field. National-level respondents also had the impression that KALRO counsels plant doctors in diagnosis, but KALRO representatives and county MoA officers said they are not in contact about plant health problems that county MoA officers are unable to diagnose.

There are currently no official early warning processes in place to identify pest and disease outbreaks through either government or PW-K sources. County officers in Nakuru briefly discussed the existence of an early warning team; a respondent from KEPHIS said, "The mission of the team is trying to identify problems that are coming up and deliberate on how such issues can be handled or managed." It was unclear how county- or national-level MoA officers identified critical information such as the fall army worm outbreak, or how POMS data is considered in doing so. One county officer from Trans Nzoia described the process of communicating field information from the ground up:

> The information flows from the farmer, to the plant doctor, to the cluster coordinator, subcounty agricultural officer, then to the county director of agriculture [CDA]. Normally the SCAOs and the CDA have monthly meetings where they discuss issues relating to all the projects that are being implemented in the county and how they are being run, and basically discuss all the challenges that the ministry faces within the county and how to resolve such issues.

Some officers also said there were challenges in relaying information through this chain; for example, one officer said, "We reported the issue to the county government and asked for assistance with chemicals, but so far nothing has been done." It also seems that, in the case of fall army worm, people in different roles and counties were simultaneously going about processes to try to find solutions, when ultimately a proven solution came from the national-level government.

B. How does the process by which PW-K is implemented influence the effectiveness of the program?

The process through which PW-K is implemented improves knowledge at multiple levels through accessible diagnosis for farmers; improved training for extension officers; and data collection to help understand where diagnosis could be improved in the short term, and where the system should address problems in the long term. In terms of program implementation at the local level, plant clinics play a key role connecting with potential beneficiaries. While there is room to improve the levels of awareness of plant clinics and their usage, the observed levels are not unusual for a demand-driven program like Plantwise, where only farmers in need of advice visit plant clinics.

Plant Doctor Assessment

A critical part of the successful implementation of PW is the training of agricultural extension agents to be plant doctors. Investing in the skills of extension agents should expand knowledge availability. Part of the evaluation was therefore to test whether the PW-K training had a significant impact on plant health knowledge. This was accomplished through a plant doctor assessment (PDA). Plant doctor training was highly regarded by every stakeholder we interviewed. Extension agents who received training learned essential information about maintaining plant health that others were not aware of (e.g., “we learned that we have an active ingredient in chemicals”).

To empirically investigate whether the PW-K training expanded participants’ plant health knowledge, we assessed all plant doctors and a group of comparable agricultural extension agents (AEAs) not involved in PW at three points in time: 2014, 2015, and 2017. In 2014, we conducted the assessment with three types of officers:

(G1) AEAs: A group of comparable extension officers who had never received PW training and serve as a control group.

(G2) Pre-2014 PDs: Extension officers who had already been trained by PW by the time we did the PDA in 2014.

(G3) 2014 PDs: Officers who had been selected to become plant doctors for the 2014 new plant clinics but had not received PW training at the time of the assessment in 2014 (a baseline for this group).

For the 2015 PDA, we added a fourth group of extension officers to the pool of 2014 participants:

(G4) 2015 PDs: Officers who were selected to become plant doctors for the new 2015 plant clinics but had not received PW training at the time of the assessment in 2015. Most of the officers from this group entered the sample for the first time in 2015.

In the 2017 PDA, we conducted the assessment for all four groups.[18] Note that officers from groups G1, G2, and G3 were assessed three times during the study (2014, 2015, and 2017). Most of the officers from the G4 group entered the sample for the first time in 2015, which means they were assessed twice (2015 and 2017).

[18] In 2016, we added a fifth group of officers (group G5). This group contained those selected to become plant doctors for the 30 clinics that planned to start in the second part of 2017, when the impact assessment was over. This group of officers was trained in 2016, right after they took the PDA. They took the PDA for a second time in 2017. This group of officers was included in the analysis in order to separate the effect of being trained from the effect of running clinics (by the time of the 2017 assessment, this group had not yet run plant clinics, despite having been trained in 2016). However, for reasons explained below, issues with the grading of the test prevent us from properly analyzing the results for this group.

In the baseline report (AIR, 2015), we showed that all groups of extension agents were very similar in most observable dimensions, including age, gender, deployment, education level, years of experience, non-PW training, consultation practices, and most of the designation and specialization categories. We also showed that the groups were different in terms of two variables. First, relative to AEAs (group G1) and current plant doctors (group G2), the group of 2014 plant doctors (group G3) was more likely to have Senior Associate Agricultural Officers (SAAOs). Second, agents with a horticulture training were more likely to be plant doctors trained before 2014 (group G2) than those with general agricultural training. Overall, we argued that the matched group of AEAs is an effective comparison group for the different groups of plant doctors.

Estimation Methodology

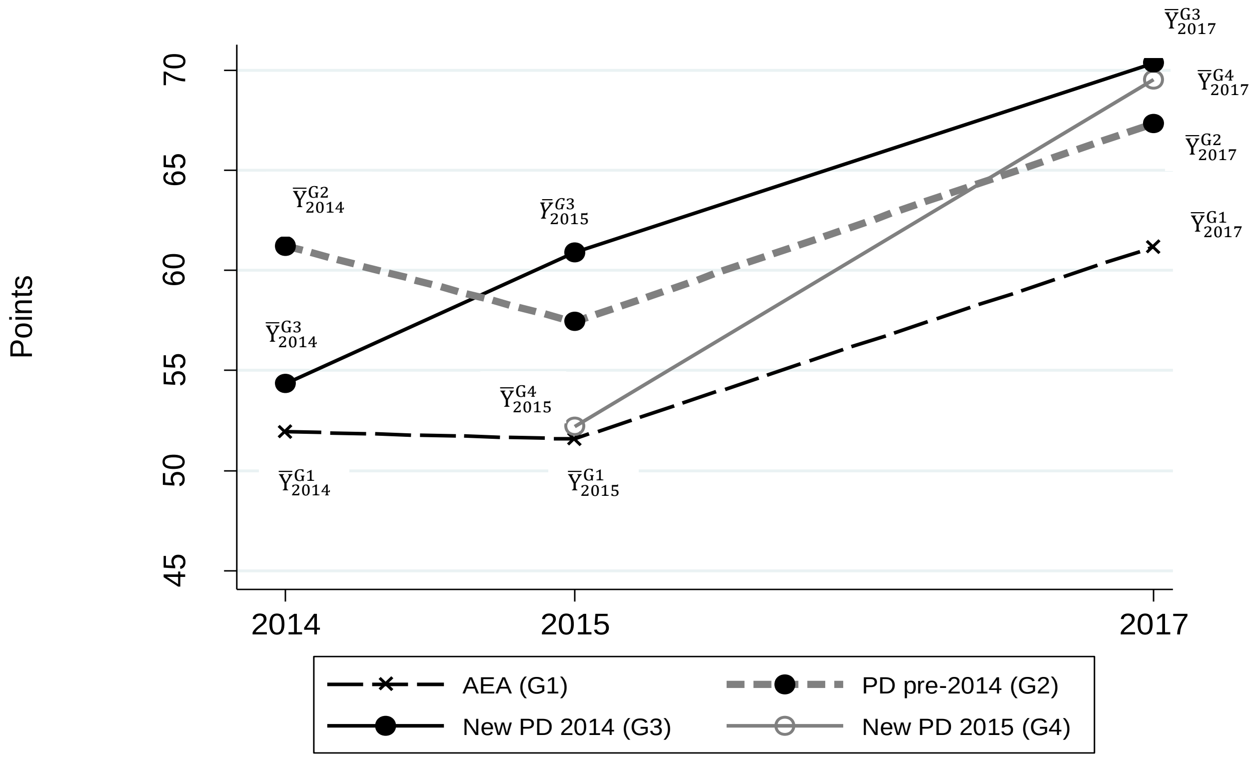
Collecting longitudinal data from training beneficiaries and a comparable group allows us to estimate the effects of PW training on plant health knowledge and recommendations. For this analysis, we used a difference in differences (DD) strategy to estimate the causal effect of PW-K trainings on plant health knowledge for those officers who were trained in 2014 and 2015 for the first time (groups G3 and G4, respectively).[19] The key insight from the DD estimation is that the effect of PW-K trainings can be estimated by comparing the average change over time in test scores for treatment groups G3 and G4, compared to the average change over time for the control group (group G1). The estimated impacts of PW training on plant health knowledge are unbiased as long as there are no differential unobserved time-varying characteristics between the treatment and control groups. This assumption is known as the parallel trend assumption in program evaluation literature. We matched AEAs and PDs from all groups to ensure that all plant doctors were as similar as possible to AEAs on key observable characteristics to increase the likelihood that assessment trends were similar for the different types of extension officers.

Figure 4.1 presents the DD estimate of the impact of PW-K training on plant health knowledge. Each line in the graph represents the trend in average scores at three points in time between 2014 and 2017 for all four groups. The black long-dashed line represents the average scores for the group of comparable AEAs not involved in PW (group G1). The gray short-dashed line shows the average scores for the PDs who received PW training before 2014 (group G2). The black solid line shows the trends in average scores for plant doctors who were trained in 2014 for the

[19] Note that we are not able to estimate the impacts of trainings for group G1 using a DD strategy because we do not have a baseline assessment for this group. We present and discuss the results for this group compared to the control group results at each point in time.

first time (group G3). Lastly, the thin gray solid line shows the mean scores for those PDs who were trained for the first time in 2015 (group G4).

Figure 4.1 Plant Doctor Assessment Results 2014–2017



Note that the 2015 DD estimate for the plant doctors trained in 2014 (group G3) can be obtained by subtracting the vertical distances between the average scores of that group and the AEAs at each point in time. That is:

$$DD_{2015}^{G3} = \{\bar{Y}_{2015}^{G3} - \bar{Y}_{2015}^{G1}\} - \{\bar{Y}_{2014}^{G3} - \bar{Y}_{2014}^{G1}\} \qquad (1)$$

Similarly, the 2017 DD estimate for the plant doctors trained in 2014 (group G3) can be obtained as:

$$DD_{2017}^{G3} = \{\bar{Y}_{2017}^{G3} - \bar{Y}_{2017}^{G1}\} - \{\bar{Y}_{2014}^{G3} - \bar{Y}_{2014}^{G1}\} \qquad (2)$$

Lastly, the 2017 DD estimate for the plant doctors trained in 2015 (group G4) can be estimated as:

$$DD_{2017}^{G4} = \{\bar{Y}_{2017}^{G4} - \bar{Y}_{2017}^{G1}\} - \{\bar{Y}_{2015}^{G4} - \bar{Y}_{2015}^{G1}\} \qquad (3)$$

The points in the graph show that AEAs (group G1) improved their score in 2017, relative to the two previous rounds.[20] The group of plant doctors who were trained for the first time after the 2014 PDA (group G3) exhibited an upward trend in scores over time. Given that these plant doctors had not received any training when they took the 2014 assessment but had received PW-K training before they took the 2015 assessment, this upward trend—and its difference from the AEA score trend—can be interpreted as the effect of the trainings on plant health knowledge. The results for the group of plant doctors who were trained in 2015 for the first time (group G4) closely follow the results of the 2014 plant doctors (group G3). The results for the G4 officers show that there was no statistically significant difference between them and the AEAs prior to receiving any PW training in 2015, but that a statistically significant difference was evident between the two groups in 2017.

Results

The overall average score for the multiple-choice section of the PDA in 2017 was 37.2 out of 50 points, and 29.2 out of 50 for the short-answer section. The total average score for the assessment in 2017 was 66.5 out of 100 points, with a standard deviation of 10 points. The assessment includes easy and challenging questions, ensuring sufficient variation in scores to identify the effects of the plant doctor training.

In Table 4.1, we present the DD results that correspond to equations 1 to 3.[21] Note that each cell in any of the panels shows the average score for a given group of extension officers at a specific administration of the PDA. For example, the top-left cell in Panel 1 is equivalent to $\bar{Y}_{2014}^{G1}$, and the top-middle cell to $\bar{Y}_{2014}^{G3}$. In turn, the center-left and center-middle cells are equivalent to $\bar{Y}_{2015}^{G1}$ and $\bar{Y}_{2015}^{G3}$, respectively. To calculate the DD estimate for this group (DD_{2015}^{G3}), as presented in Equation 1, we first subtract the average scores of each group of officers for a given PDA administration and place the results in the far-right column of Panel 1.

[20] All groups exhibited a large increase in average scores in 2017. This was due to grade inflation on the structured section of the test. The average score for the structured section of the test was nine points higher for officers of all groups who took the test in 2017, relative to the average in 2014 and 2015. The average score for the multiple-choice section was constant over time.

[21] For ease of exposition, the results presented in this section are estimated from the averages of the PDA for each group of officers over time. In Appendix 3, we present the results from estimating linear regression models that, in addition to reproducing the same results presented in this section, allow us to control for other potential determinants of plant health knowledge and estimate the associated standard errors of the impacts.

In other words, we calculate $\{\bar{Y}^{G3}_{2014} - \bar{Y}^{G1}_{2014}\}$ and place the result in the top-right cell of Panel 1, and we calculate $\{\bar{Y}^{G3}_{2015} - \bar{Y}^{G1}_{2015}\}$ and place the result in the middle-right cell of Panel 1. The DD estimate results from subtracting these two calculated differences, $\{\bar{Y}^{G3}_{2015} - \bar{Y}^{G1}_{2015}\} - \{\bar{Y}^{G3}_{2014} - \bar{Y}^{G1}_{2014}\}$, and placing the result in the bottom-right cell of the panel. The same logic applies to estimate the DD for Panels 2 and 3.

Table 4.1 The Effects of PW Training on Plant Health Knowledge

Panel 1. 2015 DD Results for 2014-Trained PDs (G3)

	AEA $(\bar{Y}^{G1})$	2014 PD $(\bar{Y}^{G3})$	$\bar{Y}^{G3} - \bar{Y}^{G1}$
2014 PDA (Baseline)	52.7	54.4	1.7
2015 PDA (F)	51.9	60.9	9.0
F – Baseline	-0.8	6.5	DD = 7.3

Panel 2. 2017 DD Results for 2014-Trained PDs (G3)

	AEA $(\bar{Y}^{G1})$	2014 PD $(\bar{Y}^{G3})$	$\bar{Y}^{G3} - \bar{Y}^{G1}$
2014 PDA (Baseline)	52.7	54.4	1.7
2017 PDA (F)	61.6	70.4	8.8
F – Baseline	8.9	16.0	DD = 7.1

Panel 3. 2017 DD Results for 2015-Trained PDs (G4)

	AEA $(\bar{Y}^{G1})$	2015 PD $(\bar{Y}^{G4})$	$\bar{Y}^{G4} - \bar{Y}^{G1}$
2015 PDA (Baseline)	51.8	52.2	0.4
2017 PDA (F)	62.4	69.9	7.5
F – Baseline	10.6	17.7	DD = 7.1

Panel 1 shows the results for Equation 1—the DD estimate in 2015 for the plant doctors trained in 2014 (group G3). The results show that the group of AEAs (the control group or G1) scored 0.8 fewer points in 2015 relative to their score in 2014, which means that the AEAs essentially had the same score, on average, in 2014 and 2015. The results also show that at baseline (i.e., when the 2014 PDA was administered), the G3 plant doctors scored just 1.7 points above the

AEAs—an estimate that is not statistically significant. In other words, untrained officers performed equally in the test in 2014, regardless of whether they had been selected to receive training later that year. The 2015 DD results on the effect of trainings show that the plant doctors who were trained by PW-K in 2014 (group G3) scored up to 7.3 more points in 2015, relative to the control group of AEAs (group G1).[22] This estimate represents a large and statistically significant difference in plant health knowledge. Given that the average AEA scored 51.9 points in the assessment (with a standard deviation of 10), scoring 7.3 additional points is equivalent to a 0.7 standard deviation gain, which is considered a large effect in the assessment literature.

In Panel 2, we present the results for Equation 2—the DD impact of PW training for those officers who were trained in 2014 (group G3), using the scores from the 2017 PDA. The results show that PDs who were trained in 2014 scored 7.1 more points on the assessment, relative to the control group. The estimated impact is very similar to the impact estimated in 2015 (Panel 1) and provides evidence that the gap in plant health knowledge between trained and untrained officers is not changing over time. The results also show that, relative to the scores obtained on the 2014 PDA, the group of AEAs (group G1) and the group of trained plant doctors (group G3) had substantial increases in average scores in the 2017 PDA: the AEA officers (group G1) increased their average score by 8.9 points over time, while the 2014-trained plant doctors (group G3) increased their average score by 16 points. This increase in average scores was driven by a generalized increase in scores for the structured (open-ended) questions in the test. We conclude that this increase is due to grade inflation in the 2017 PDA.[23]

In Panel 3, we show the impact of PW training for plant doctors who were trained in 2015 for the first time (group G4). The results are again very similar to the DD estimates for the cohort of plant doctors trained in 2014 (group G3). In 2015 (the baseline for this analysis), officers who had been selected to be plant doctors scored only 0.4 points higher than the control group (group G1). That means that the treatment and control groups had the same level of plant health knowledge at baseline, as expected. The DD estimate shows that officers trained in 2015 (group G4) scored 7.1 points higher than the control group in the 2017 PDA. This estimated impact is very similar and not statistically different from the DD estimate of the plant doctors trained in 2014 (group G3).

[22] These two DD estimates are not statistically different (p-value = 0.75), which means there is no evidence that the impacts of the training decrease over time.

[23] If there had been an increase in overall plant health knowledge, we would have observed an increase in the average score of the multiple-choice section of the test as well. However, the mean of the multiple-choice section is the same as averages observed in previous rounds of the PDA.

Lastly, in Appendix 3, we estimate the effects of PW-K trainings for those plant doctors who had already been trained by PW-K at the time of the first PDA in 2014. The results show that in 2014, this group of plant doctors (group G2) scored between eight and nine points higher than the AEAs (group G1). Although we do not have assessment data for these two groups before the G2 plant doctors were trained the first time (i.e., we do not have baseline data that allows us to implement a DD strategy for group G2), it is encouraging that these estimates are similar to the estimates of the effect of training for the 2014 plant doctors (group G3) in 2015 and in 2017. The results also show that the G2 plant doctors outperformed the AEAs (group G1) by approximately five test points in 2015 and in 2017, which is equivalent to 0.5 standard deviations—again a significant difference.

Overall, we interpret the similarity of the estimated training impacts on plant health knowledge across years and different cohorts of plant doctors (groups G3 and G4) as strong evidence that trainings provided by PW-K produce a large and significant effect on plant health knowledge.[24]

Qualitative data indicated that farmers attended clinics because they trusted the doctors were educated. Farmers also said plant doctors' solutions were more useful than those provided elsewhere when there was "an emergency." One farmer said, "The diseases are not there but when you chat with [the doctors], you get to learn of new things, and you get information before the need arises." Plant doctors also felt more educated themselves; for example, one plant doctor stated:

> When you go to a meeting, you have good substance when with farmers, and this is Plantwise. You have substance when talking. When we were doing coffee the early days, farmers were not seeing you as very technical, but after the training we could go and emphasize a problem using the technique of Plantwise and they acknowledge that.

Overall, we continue to see important knowledge gains through the PDA. This finding aligns with baseline information that indicated that AEAs regard their PD peers as having more

[24] As mentioned above, we also administered the PDA in 2016 to the group of officers who were selected to run the last set of 30 plant clinics that started at the end of 2017 (group G5). We conducted the assessment of this group of officers before they were trained in 2016 (their baseline) and repeated the assessment with them in 2017 (their follow-up), one year after being trained. For logistical reasons, the baseline and follow-up assessments for the group of 2016 plant doctors were graded at the same time in November 2017. Analysis shows that the average scores in 2017 were much higher for the structured section of the test for all groups of officers, including the baseline scores for the 2016 cohort of plant doctors. The fact that the baseline of the comparison group was graded in 2015, when graders were not inflated mechanically, implies that the comparison group exhibited a large increase in scores between 2015 and 2017. The increase in scores for the 2016 cohort is more modest. Thus, the estimated DD impact of training on the structured section of the test is negative for G5, which would imply that receiving PW training decreases plant health knowledge.

knowledge, as well as access to more knowledge. Endline qualitative data indicated some areas for improvement, including that plant doctors could benefit from additional assistance from officers of agriculture or research agencies, and from paper materials, trackable targets for farmer outreach, and help in understanding that farmers may not be able to afford some of the suggested solutions.

Expanded Knowledge Availability and Use

Plant doctors described the process by which they identified a pest or disease in a detailed way, indicating that training had encouraged them to make a critical assessment in order to advise farmers. One plant doctor reported that they give more general recommendations to most farmers, and more specific recommendations to farmers in situations they are more familiar with:

> For recommendations, we start with stressing pest management so you can prevent diseases. After giving all those, it is the farmers' option to see which to follow. For example, there are new diseases of maize [for which] we do not have chemicals. But if … you happen to know the farmer, you stress on rotation, or [advise to] stop planting that [if] the previous plot had a problem. So, you find that situation where you stress a particular recommendation … otherwise we give all recommendations [for the] future so that you know how to handle when you plant.

Another plant doctor described a similar process: "When a farmer comes to the clinic, you don't start recommending, you probe. Did you use another chemical, did you use manure? You will ask them, do you have goats, do you have sheep, how many are they? So, you [understand they are able to] accumulate the manure." Plant doctors also mentioned that they considered the level of infestation and stage of growth when giving recommendations.

Farmers said they primarily got their information on plant health from the clinics, and that getting information without the clinics would be difficult. Plant doctors are likely to use the Knowledge Bank or fact sheets for information. When the farmers were asked where they would go for advice if plant clinics did not exist, very few said they would consult the MoA. One person said that the Office of Agriculture was not the best alternative because of its distance from farmers: "We can go to the Office of Agriculture, but you see when they come to our land it makes it easy, so that when they are asked about some area, they can be able to say what [the] crop is and how the crop is." Another farmer described the situation before the plant clinics existed: "When we didn't have the doctors or the clinics, many farmers, for example my case—I was afraid of doing farming, I was afraid of planting watermelons, green grams. I didn't know how to deal with the diseases, [so] I would just abandon them."

Quantitative results suggest that treatment farmers were more likely to report receiving information on a variety of agricultural practices (Table 4.2). Specifically, treatment farmers were more likely than control farmers to have received advice on new seed varieties (4 percentage points more likely), pest control (6 percentage points more likely), postharvest technologies (1 percentage point more likely), and value addition/agroprocessing (1 percentage point more likely). Interestingly, the largest impacts are observed for receiving advice on new seed varieties and pest control, which are common recommendations given by plant doctors. These results take into account information from any source, including plant clinics.

Table 4.2 Receipt of Crop Information (Single Difference)

Dependent variable Farmer receives advice on…	ITT impact (1)	Control mean (2)	Treatment mean (3)	N (4)
New seed varieties	0.04** (2.44)	0.21	0.25	2,550
Pest control	0.06*** (3.57)	0.18	0.24	2,550
Fertilizer use	0.02 (1.56)	0.17	0.19	2,550
Irrigation	0.00 (0.36)	0.03	0.04	2,550
Marketing or crop sales	-0.01 (-0.66)	0.04	0.04	2,550
Postharvest technologies	0.01** (2.43)	0.02	0.03	2,550
Value addition/agroprocessing	0.01* (1.90)	0.01	0.01	2,550
Farmer received useful information on…				
New seed variety	0.04** (2.35)	0.94	0.98	595
Pest control	0.01 (0.81)	0.97	0.98	555
Fertilizer use	-0.01 (-0.74)	0.99	0.97	475

Note. Robust *t* statistics clustered at the clinic level in parentheses. All estimations control for lottery fixed effects. Impact estimates are 36-month impact estimates. Means are 36-month means. ITT=intention to treat. *p < .10; **p < .05; ***p < .01.

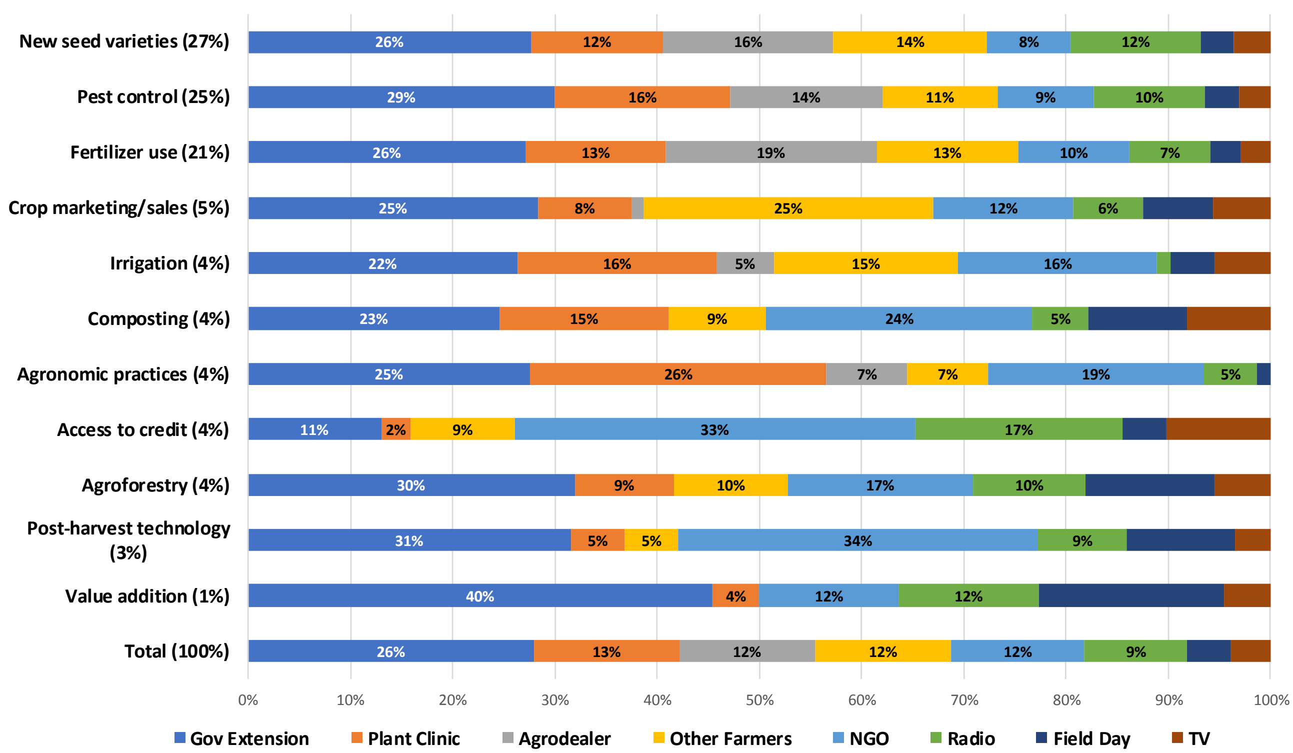
Farmers received information from multiple sources, in addition to plant clinics. Results show that getting information from government extension officers was the most important information source for farmers, 28% of those in the treatment group reported at endline that they had received agricultural information from that source relative to 19% of the control group. It is possible that farmers do not make a distinction between the government extension officers and the plant clinics as technically officers who serve at plant clinics are government extension officers. That may explain why the proportion of farmers that report getting advice from a government extension service is higher in treatment than in control areas. The second most important information source was other farmers, from which 11% of farmers in the treatment group reported receiving information compared to 15% in the control group. The next most important source was plant clinics, from where 15% of farmers in the treatment group got information. Interestingly, 7% of farmers in the control group also reported receiving information from plant clinics. The next most important sources of information, in descending order, were NGO (12%), agrodealers (12%), and the radio (9%). Table 4.3 presents the proportion of farm households that reported receiving agricultural information from a given source for any topic (see Figure 4.2 for the list of topics). The table presents information only for those who receive information from at least one source. Note that 88% of farmers in the endline reported receiving information from at least one source. The table also compares the use of information sources by treatment condition.

Table 4.3 Crop Information Sources for Any Agricultural Topic

Source	Control	Treatment	Total
Government agricultural extension service	0.19	0.28	0.25
Other farmer (neighbor/relative)	0.15	0.11	0.12
Plant clinic	0.07	0.15	0.12
Non-governmental organization	0.16	0.10	0.12
Agrodealer	0.11	0.12	0.12
Electronic media: radio	0.10	0.09	0.9
Field days, shows, fairs, or field school	0.05	0.03	0.4
Electronic media: TV	0.04	0.03	0.3
Agricultural co-op or farmers’ association	0.02	0.02	0.02
Private agricultural extension service	0.02	0.02	0.02
Commodity-based extension services	0.03	0.01	0.02
Village agricultural extension meeting	0.02	0.01	0.01
Agricultural training centers	0.01	0.01	0.01
Agricultural extension course	0	0.01	0.01
Champion/contact farmer	0.02	0	0.01
Electronic media: internet	0	0	0

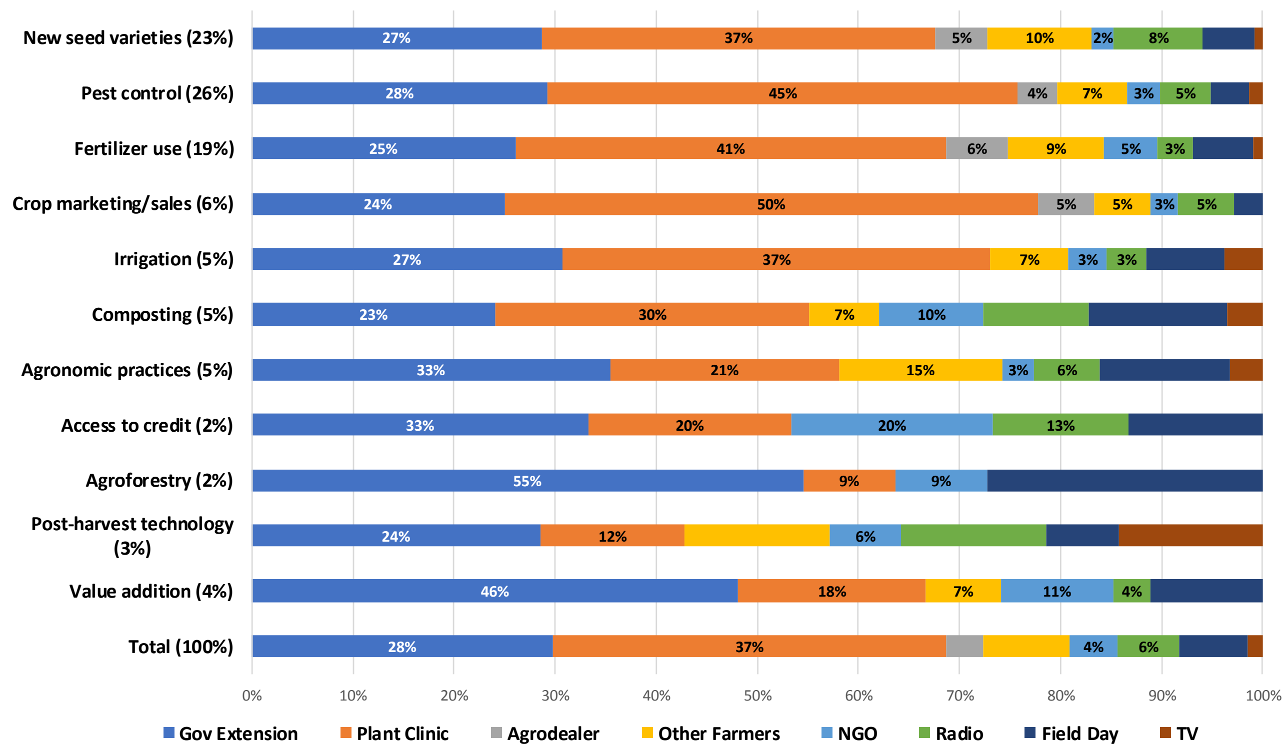
In general, farmers use the same type of sources regardless of the type of advice they request. However, in some cases we observe that there are some specific sources that are more relevant for a given topic. The most relevant information source for almost all type of topics is government extension agents. Plant clinics are also important for some key topics. For instance, plant clinics are the most relevant source when farmers need advice on agronomic practices and the second most relevant source for pest control. Agrodealers are also a common source for those seeking advice on fertilizer use, seed varieties, and pest control. Fellow farmers are also an important source of information specially for topics on crop sales, irrigation, seed varieties, fertilizer, and pest control. The next sources that are relevant are NGOs and the radio. NGOs seem to be a common information source for advice on access to credit, composting, agroforestry, and irrigation. Lastly, the data show that farmers get more or less the same rate of advice on the radio for most topics. Figure 4.2 presents the distribution of information sources by the type of advice requested.

Figure 4.2 Distribution of Crop Information Sources by Topic of Advice Received



We also explored the distribution of crop information sources for those farmers who use clinics and for those who live in treatment areas (Figure 4.3). Clinic users say that the most important source of information for almost all topics is the plant clinics, with very little information being obtained from other sources such as agrodealers and fellow farmers. These results provide evidence that plant clinics are providing information beyond plant health issues.

Figure 4.3 Distribution of Crop Information Sources by Topic (Clinic Users Only)



In additional analyses, we examined farmers' reported reasons for increased crop yields. We found that farmers identified improvement in pest information as a main reason for increases in crops yields, both when looking at all treatment farmers (Table 4.21) and when looking at maize producers (Appendix 1, Table A1.26).

Plant Clinic Use

Increasing farmers' awareness and use of plant clinics is a key aspect of PW-K implementation. We investigated clinic awareness and use by treatment condition, using farm-level data. Table 4.4 shows the number and proportion of farms that reported being aware of the existence of a plant clinic and attending a clinic. The top panel of the table reproduces the figures from the midline report. The bottom panel presents information on clinic awareness and use based on both midline and endline data. We considered a farm household to be aware of plant clinics if it reported being aware at least once in the 2015 and 2017 rounds of data collection. Compared to midline, a higher fraction of treatment farms were aware of plant clinics at endline, and treatment farms were more likely to have visited a clinic. At midline, 38% of treatment farms reported being aware of clinics, compared to 53% at endline.[25] In the control group, the proportion of farms that were aware of plant clinics also increased from midline to endline, from 10% to 30%. In terms of attendance, the proportion of treatment farms that reported attending a clinic at least once by endline was 34%, compared to 30% at midline. The proportion of control farms that reported attending a plant clinic at least once increased from 9% to 24% between midline and endline.[26]

Table 4.4 Farmers' Awareness and Attendance of Plant Clinics in 2015 and 2017

Midline (2015)	Awareness		Attendance	
	Treatment	Control	Treatment	Control
No	532	1502	229	154
	62%	90%	70%	91%
Yes	326	169	97	15
	38%	10%	30%	9%
Total	858	1671	326	169
	100%	100%	100%	100%

Endline (2017)	Awareness		Attendance	
	Treatment	Control	Treatment	Control
No	794	598	604	190
	47%	70%	66%	76%
Yes	909	249	305	59

[25] If we used only the responses from the 2017 round of data collection, we found that the proportions of treatment and control farmers who were aware of plant clinics were 45% and 23%, respectively. This means that the awareness proportions reported in Table 4.2 are approximately 6 percentage points higher when we combine the responses from the midline and endline rounds of data. This difference is due to the reporting of 148 farms, which reported that they were aware of plant clinics at midline but reported that they were not aware of plant clinics at endline. The main survey respondent may have changed between waves of data collection, which may partly explain the reported difference. Misreporting awareness in any of the rounds is also possible. The 148 cases represent just 6% of the households surveyed at endline.

[26] Only 68 households that reported visiting a plant clinic in the 12 months before the 2015 midline data collection did not report visiting a plant clinic in the 12 months before the 2017 endline data collection. Only 34 households reported visiting a plant clinic in both waves of data collection. Most of the growth in clinic attendance was driven by households that had not attended a clinic at midline but had attended at least one clinic by endline.

	53%	30%	34%	24%
Total	1703	847	909	249
	100%	100%	100%	100%

We can derive some key insights from the data on clinic awareness and use. First, while there is room for increasing farmer awareness of plant clinics, the plant clinic attendance levels are not unusual for a public-demand-driven program like PW. Only a fraction of the potential beneficiaries who are offered the intervention are expected to directly take part in program activities because, just like human health clinics, beneficiaries are not expected to come on a regular basis.

Second, as discussed above, we observed an increase in the proportion of control farmers—although not so much in the number of farmers—who reported attending plant clinics at least once in any of the two follow-up rounds of data collection. Fifteen farmers reported attending a plant clinic session in the 12 months before the 2015 survey, and 59 farmers reported attending a plant clinic in any of the survey rounds in 2015 and 2017.[27] Control farmers' use of these clinics reflects well on the services they provide, given that these farmers have to travel a longer distance than farmers from treatment areas to attend a clinic (an average of 4.7 km for control farmers, compared to an average of 1.6 km for treatment farmers). In principle, ITT program estimates could be biased by control farmers attending and benefiting from plant clinics (see Section 3A), as these estimates compare farmer outcomes in the treatment group to farmer outcomes in the control group. However, concerns about biasing ITT estimates are limited, given that the number of control farmers attending clinics is low, and that PW benefits go beyond plant clinic attendance.

Characterizing Users of Plant Clinics

We also used data from the endline survey to characterize the profile of plant clinic users. We present this information in Table 4.5. In Panel 1, we compare the observable characteristics of those who attended at least one clinic session in the 12 months before the 2017 survey with the observable characteristics of those who did not attend a clinic. Compared to non-users in our sample, those who attended plant clinics were more likely to be male, more likely to have more years of education, to be members of farmer associations and other groups, to have larger farms, and to have a more diverse set of assets, and more likely to live closer to clinics.

[27] 50 of the 59 farmers attended a clinic in the 12 months before the 2017 survey.

On average, plant clinic users in our sample were middle-aged men with very low educational attainment[28] (mostly some years of secondary education), who were slightly better-off in terms of their economic characteristics. However, it is important to note that our sample only includes smallholder farmers who were willing to receive extension advice if given, and who live close to plant clinics. We do not know how clinic users in our sample compare to the average farmer in the Kenyan population.

Table 4.5 Descriptive Characteristics of Plant Clinic Users and Non-Users

	Panel 1. All households		Panel 2. Clinic users	
	Users (n=296)	Non-users (n=2254)	Control (n=50)	Treatment (n=246)
Gender of HH head (male=1)	0.89***	0.78	0.86	0.89
Age (years)	48.99	49.89	48.48	49.09
Primary education or less	0.47**	0.54	0.54	0.46
Secondary education	0.39	0.37	0.34	0.4
Tertiary education	0.14***	0.09	0.12	0.14
Farming experience (years)	22.75	24.63	17.06	19.5
Non-farm income (KSH)	5888	6619	5110	6046
Farmer association	0.17***	0.08	0.16	0.17
Group membership	0.80***	0.70	0.76	0.81
Land holdings (acres)	4.15*	3.43	3.94	4.19
Livestock holdings (number)	18.3	15.3	24.0	17.2
Asset index	0.25**	-0.03	0.51	0.19
Distance to closest clinic (km)	2.12**	2.53	4.7***	1.6

Note. Table compares household-level characteristics by clinic users versus non-users (Panel 1) and by treatment and control users (Panel 2).
*p < .10; **p < .05; ***p < .01.

In line with the quantitative data, some plant doctors reported that male farmers attended clinics more frequently than female farmers. Male plant doctors reported that this happens because men have more free time than women. Female plant doctors indicated that women typically have to consult their husbands (where applicable) about spending money on a solution to an agricultural challenge. A desk officer described the issue of women's attendance as follows:

> Most of the farming activities are implemented by female farmers, but if we look at attendance … we see that there are more men attending than women, and so we are not addressing the right people. We tried to find out why, and some of

[28] Compared to the general population in Kenya. However, as mentioned earlier, clinic users in our sample are slightly more educated than non-users in our sample.

> the reasons we found are [that] men mostly go to the market in the evenings and so the women would prefer to send their husbands to take such issues to the clinics. Then there is the gender roles issue, it is the man who is supposed to purchase inputs and not women.

The desk officer also said they tried to “take some measure to encourage more women to participate,” such as bringing clinics closer to areas where women hold groups; however, the desk officer did not say whether he perceived any changes as a result of the efforts. One plant doctor also said, “The woman does all the work, but when it comes to selling, the woman has no say in decision making.” Many respondents echoed this sentiment. Finally, women also seem to have more challenges with transport because they cannot or do not ride motorbikes.

In Panel 2, we compare the characteristics of clinic users by their treatment status. Control and treatment farmers who attended plant clinics were very similar, on average. With the exception of distance to the closest clinic, none of the differences in characteristics were statistically significant. As expected, the average distance to the closest plant clinic was 4.7 km for farmers in the control group, compared to just 1.6 km for farmers in the treatment group. The similarity of treatment and control clinic users is not surprising, given that farmers from both groups in our sample were very similar in most dimensions and the treatment status was randomly assigned.

We also used quantitative and qualitative data to investigate the reasons that may be driving farmers to attend (or not attend) plant clinics. In the qualitative data, the most frequently cited reason for attending clinics was that they were free. Qualitative data indicate that farmers who attended clinics found them useful, not only for understanding how to treat pests and diseases, but also for discussing general farming issues or reading about related topics. One farmer said, “At times when I notice a disease on a tree, and we … cut a leaf and bring it to [these doctors], they can give us a prescription. In other occasions, if I see worms and I don’t know what spray to use, they will tell me what chemical to use.” The data from the quantitative survey supports the claim that farmers go to clinics for general advice, even if they do not have a pest problem. While 63% of the farmers who attended plant clinics stated that the primary purpose of a plant clinic is to diagnose plant health problems, 35% said that plant doctors provided advice on crop production in general. The quantitative data also show that farmers who experienced any damage due to insects, fungus, disease, or other pests in any of their crops were more likely to have visited a plant clinic in the 12 months before the 2017 survey. As shown in Table 4.6, farmers who attended plant clinics were 5 percentage points (5pp = 44% - 39%) more likely to report that they experienced a plant health issue with any of their crops. However, the data do not show that farmers who attended plant clinics reported larger crop losses than farmers who did not attend. Both groups reported that whenever they had a plant health problem, close to 40% of their harvest was affected.

Table 4.6 Propensity of Attending a Plant Clinic by Crop Damage Status

Damage to any crop in the last 12 months	Attended plant clinic in the last 12 months	
	No	Yes
No	1,377 (61%)	165 (56%)
Yes	877 (39%)	131 (44%)
Total	2,254 (100%)	296 (100%)

The main reason farmers offered for not attending clinics was a lack of need. A plant doctor suggested that some farmers did not visit clinics because they were further away than the agrodealers' shops, and the farmers preferred not to waste time traveling. One cluster coordinator said about clinic attendance: "I think the farmers expect the extension officers to visit them at home, because any time we visit them in their farms we must always come across some issues. Unless it is a very pressing issue, the farmers might not come to the clinics at all."

A female farmer indicated that other farmers might not attend clinics because they felt they could get the information from neighbors who did attend. Farmers who did not attend clinics asked neighbors who attended about problems, saying, "You are the ones who go to that clinic." This is consistent with the data presented in Table 4.3 on crop information sources, which shows that neighbors are the second most common source of agricultural information, after government extension officers. This highlights the importance of extending the assessment of program implementation beyond measuring plant clinic attendance, as there are other ways in which PW can potentially benefit farmers who live in a plant clinic catchment area.

Farmers' Use of Solutions from Plant Doctors

Qualitative data indicated that farmers may not consistently implement the prescribed remedy for a pest or disease, primarily because of the associated costs. Farmers said they faced challenges with the costs of prescribed remedies, despite some plant doctors saying they tried to tailor recommendations to farmers based on cost. One plant doctor said that because small-scale farmers have trouble affording expensive solutions, "You recommend something he can afford." Some farmers said they did not tell the plant doctor that they needed a less-expensive alternative remedy because they were embarrassed. One farmer said:

> It happens. They can give you an expensive prescription—more than you can afford—so you look for an alternative. You tell them to give you a different prescription. I tell him I have 1 acre, and they can tell me to use a chemical [that costs] 180 shillings for each spray can. So, I get a different alternative. So not all the time you do what they tell you to do.

One farmer mentioned that chemical companies were more helpful with cost-related challenges, as farmers were able to negotiate paying for the chemical after their yield: “The chemical companies come to the groups, so when they come to sell the chemicals we can explain to them that we don’t have enough money but we will pay after we sell our produce.” Finally, according to treatment farmers, their neighbors were likely to uproot a single plant with a disease and forget about it, rather than report it to plant doctors. This aligns with the finding presented earlier that farmers only attend clinics once a problem has become more serious.

C. Does PW-K improve the wellbeing of farmers?

The evidence suggests that, because of their training, the plant doctors promote lasting improvements in knowledge. PW-K aims for plant doctors to use this knowledge to promote crop protection, which allows households to avoid and reduce damage from pests and disease. In the short term, we expected the program to achieve intermediate outcomes at the farm and institution level that would facilitate achievement of the final impacts. This section discusses intermediate changes in farmers’ use of cultural practices and inputs, and in production diversity, as well as how program activities and outputs affected intermediate institutional outcomes.

Production Descriptive Statistics

Before discussing the estimated program impacts on intermediate and final outcomes, we describe some key features related to crop production in our sample. In particular, we provide evidence for why we look at the impacts on some specific crops in our sample and not others. In Table 4.7, we present the distribution of crop production by county and show that there is significant regional variation in agricultural production. The table includes the most common annual and perennial crops that are produced by at least 5% of farmers in our sample. In general, maize were the most common crops across all counties. There is also a large degree of regional variation in terms of crops cultivated. In Embu, for example, crops that were also important include coffee and bananas; peas and mangoes in Machakos; potatoes in Elgeyo and Nakuru; coffee, bananas, and napier grass in Nyeri; and bananas and coffee Tharaka-Nithi. West Pokot showed the least crop diversification. Overall, the table show that, after maize and beans, the most common crops produced are bananas, avocados, potatoes, coffee, and kale.

Table 4.7 Proportion of Households Producing Crop, by County

Most Common Annual Crops						
County	Maize	Beans	Potatoes	Kale	Sugar Cane	P. Peas
Bungoma	99.4	73.8	0.6	19.2	14.5	-
Elgeyo	82.1	39.9	83.3	38.0	1.1	-
Embu	96.2	76.2	10.8	7.6	19.5	-
Kajiado	67.4	46.5	6.9	12.2	-	1.2
Kiambu	77.5	62.9	39.1	25.8	3.3	-
Kirinyaga	75.5	50.9	5.4	1.2	2.4	0.6
Machakos	100	88.8	1.2	8.9	2.4	60.9
Nakuru	98.8	84.4	66.3	46.9	6.3	1.3
Narok	90.6	53.6	20.4	8.7	0.8	-
Nyeri	96	74.3	14	3.5	3.5	-
Tharaka Nithi	75.8	53.6	4.4	4.4	16.5	8.5
Trans Nzoia	97.7	70.1	5.6	18.6	4	0.6
West Pokot	98.9	70.1	22	22.6	3.4	-
All Counties	**87.8**	**62.8**	**22.4**	**16.5**	**5.8**	**5.1**
Most Common Perennial Crops						
County	**Bananas**	**Avocados**	**Coffee**	**Napier**	**Woodlot**	**Mangoes**
Bungoma	54.7	41.3	24.4	14.0	33.7	20.9
Elgeyo	12.5	23.2	-	6.8	5.3	-
Embu	85.9	37.8	74.1	43.8	29.2	15.7
Kajiado	2.0	0.4	0.4	0.8	2.0	0.4
Kiambu	21.2	29.1	-	23.2	-	10.6
Kirinyaga	32.9	10.2	3.6	3.6	-	13.2
Machakos	21.3	23.7	14.2	3.0	20.7	47.9
Nakuru	39.4	50.6	6.3	16.9	2.5	3.8
Narok	2.6	1.5	0.8	3.4	0.8	0.4
Nyeri	48.5	18.7	74.3	40.4	-	10.5
Tharaka Nithi	58.1	24.2	39.5	31.5	18.1	16.5
Trans Nzoia	38.4	29.4	0.6	14.7	14.1	1.1
West Pokot	27.7	28.8	5.6	9.6	22.0	1.7
All Counties	**32.5**	**22.9**	**18.0**	**15.6**	**11.0**	**10.0**

Note. This table presents the proportion of households by county that produce at least one kilogram of the crop. Only crops produced by more than 5% of the farms are included.

The information provided in Table 4.7 does not consider the size of the area where the crops are produced. This is relevant for this study because we only collected detailed data on production for crops produced in areas larger than 1/32 acre. In Table 4.8, we present the average area used in the production of some selected crops and the number farms that produce the crop in areas larger than the specified threshold. The results show that the average area used to produce maize in our sample is 1.38 acres, which is similar to the area used to produce bananas, potatoes, and beans. Other crops such as coffee and kale, are produced in areas that are on average cultivated in 0.5 acres. In terms of the number of farms for which we collected production data, we see that we have relatively good sample sizes for maize (n=2,136) and beans (n=1,479) to estimate program impacts. However, for the other crops reported in Table 4.8, the number of observations is rather low to detect program impacts.[29] Thus, when estimating program impacts, we focus our attention on all crops aggregated at their plant cycle level (annual and perennials) as well as maize. However, for descriptive purposes, we also present results for some key crops such as beans, potatoes, coffee, and kale, which are crops that are commonly brought to plant clinics by farmers.[30]

Table 4.8 Average Area of Most Common Crops (Acres)

	Control	Treatment	Total
Maize	1.40 [700]	1.38 [1,436]	1.38 [2,136]
Beans	1.29 [523]	1.18 [956]	1.22 [1,479]
Potatoes	1.21 [134]	1.27 [278]	1.25 [412]
Coffee	0.46 [140]	0.54 [251]	0.51 [391]
Bananas	1.02 [105]	1.15 [208]	1.11 [313]
Avocados	1.80 [36]	2.4 [91]	2.2 [127]
Kale	0.30 [40]	0.62 [56]	0.49 [96]

Note. The first row for each crop is the average area of land used for production and the second row the number of farms with areas larger than 1/32 acre.

[29] Recall that in our power calculation we determine that we needed a sample of 2800 observation to find an impact of 0.2 standard deviations. For a sample size as the one we have for maize, the minimum detectable effect size (MDES) is 0.25 SD. However, the MDES for smaller samples like the ones we have for the other crops in the table is close to 0.4 SD.

[30] The most common crops brought to plant clinics in Kenya are: maize (16%), coffee (12%), kale (11%), and tomatoes (10%).

Cultural Practices and Input Use

We explore the impact of PW-K on intermediate outcomes related to farmers' use of cultural practices and inputs. All quantitative tables in this and the following sections follow a format that provides information about program impacts at 36 months, as well as 36-month statistics. Column (1) in each of these tables shows the intent-to-treat (ITT) impact of PW-K between baseline and 36 months. Columns (2) and (3) show the mean values for the treatment and control groups at 36 months. These are important in assessing the magnitude of the estimated impacts reported in Column (1). The *t* statistic, shown in parentheses under the impact estimate, is used to help determine statistical significance.

First, we present the impact estimates for all the crops produced on an area of land greater than 1/32 acre aggregated at the farm level. Then we present the impact estimates for annual and perennial crops that were produced also on areas greater than 1/32 acre. Tables 4.9–4.12 present the impact estimates for use of cultural practices and inputs. Table 4.9 includes the practices that are common to annual and perennial crops. We see no statistically significant effects on the use of fertilizer or pesticide (Table 4.9).

Table 4.9 Impacts on Cultural Practices and Input Use: All Crops (ANCOVA)

Dependent variable	ITT impact (1)	Control mean (2)	Treatment mean (3)	N (4)
Organic fertilizer used	-0.02 (-1.07)	0.58	0.55	2,440
Inorganic fertilizer used	-0.01 (-0.32)	0.82	0.83	2,440
Pesticide used	-0.02 (-0.72)	0.50	0.49	2,440

Note. Robust *t* statistics clustered at the clinic level in parentheses. All estimations control for lottery fixed effects and outcome value at baseline. Impact estimates are 36-month impact estimates and outcome value at baseline. Means are 36-month means. ITT = intention to treat. *p < .10; **p < .05; ***p < .01.

Tables 4.10 and 4.11 present impacts for annual and perennial crops, respectively. Based on our intention-to-treat estimates, Table 4.10 shows that treatment farmers are 4 percentage points more likely than control farmers to practice crop rotation, remove volunteer crops, and remove infested or damaged material with their annual crops.[31,32] In addition, we found statistically

[31] The results are statistically significant at the 5% level except results for crop rotation, which are weakly significant at the 10% level.

[32] We investigated the reasons behind crop rotation; 49% of farmers said they do it to improve soil fertility, 27% to avoid pest or diseases, and 22% to have higher yield. There were no differences in reasons to practice crop rotation between treatment and control.

significant results at the 5% and 1% levels that farmers were more likely to apply ash and spray with chilies, respectively, for their annual crops.[33] We also found a statistically significant decrease in the likelihood that treatment farmers would practice intercropping with their annual crops. This result is surprising, since intercropping is a common pest control recommendation—especially the intercropping of maize and beans, which are the most common crops in our sample. Interestingly, among farmers who intercropped, less than 1% of the farmers said that the main reason to intercrop is to avoid crop infestation. Sixty-nine percent said they do it to make the most use of land, 18% to maximize yields and profits, and 10% to improve soil fertility. There were no differences in treatment and control in terms of the reasons to intercrop.[34]

We also found little change in the use of cultural practices or inputs for treatment farmers with regard to their perennial crops. Table 4.11 shows these results, which are not statistically significant except for the results suggesting a decrease in the use of pesticide, which we explore in more detail below.

We also looked at program impacts on cultural practices for the subsample of farmers who produce maize—a crop that 87 percent of farmers in our sample produce in areas above the 1/32-acre relevance threshold.[35] We examined impact estimates for other individual crops and present the results in Appendix 1. For the most part, the results are similar to the ones observed for maize.

Table 4.12[36] contains the results for maize producers. It shows that, as with the results for annuals in Table 4.10, treatment farmers were more likely to use a variety of cultural practices. Specifically, of farmers producing maize, treatment farmers were more likely to practice crop rotation, spray with chilies, and apply sand with their maize production.[37] Treatment farmers were less likely to intercrop with their maize—a result that can be interpreted as farmers concentrating more on maize production.

---

[33] These two impacts are concentrated on maize and bean producers. While the impacts are statistically significant, note that these practices are very rarely used, especially using chilies, which only 1% of the sample practices.

[34] We are unable to estimate the impacts of the program on adopting biological crop protection practices (use of microbials, macrobials, traps, and plant extracts) because only five farms reported any of those practices.

[35] The analysis of POMS data also shows that maize is the most common crop brought to plant clinics by farmers.

[36] Because farmers changed their production patterns over time to produce more maize, the ANCOVA estimate for maize production is not appropriate. That is, because fewer farmers produced maize at baseline, it did not make sense to control for baseline values of maize cultural practices and input use. Instead, we report the maize estimate on the basis of the single difference, comparing treatment and control only on endline results.

[37] Again, spraying with chilies and applying sand, though statistically significant, are practices used by a very low proportion of farmers.

Table 4.10 Impacts on Cultural Practices and Input Use: Annual Crops (ANCOVA)

Dependent variable	ITT impact (1)	Control mean (2)	Treatment mean (3)	N (4)
Organic fertilizer used	-0.03 (-1.64)	0.46	0.43	2,363
Inorganic fertilizer used	-0.01 (-0.41)	0.82	0.82	2,363
Pesticide used	-0.01 (-0.56)	0.42	0.41	2,363
Crop rotation	0.04* (1.96)	0.27	0.31	2,363
Used resistant variety	0.02 (1.14)	0.45	0.48	2,363
Used improved planting material	0.02 (0.94)	0.76	0.79	2,363
Used certified planting material	-0.02 (-0.72)	0.84	0.83	2,363
Removed plant residue from prior harvest	0.01 (0.64)	0.94	0.94	2,363
Planted early	0.02 (1.06)	0.74	0.76	2,363
Intercrop	-0.07*** (-2.69)	0.65	0.58	2,363
No. times crop was checked	-0.30 (-0.43)	14.26	14.19	2,363
Weeded in a timely manner	-0.00 (-0.04)	0.98	0.97	2,363
Removed volunteer crops	0.04** (2.07)	0.66	0.71	2,363
Removed infested or damaged material	0.04** (2.00)	0.62	0.65	2,363
Applied ash	0.03** (2.46)	0.05	0.08	2,363
Sprayed with chilies	0.01*** (3.00)	0.01	0.02	2,363
Staked	0.01 (0.64)	0.07	0.08	2,363
Applied sand	0.01 (1.56)	0.06	0.07	2,363
Used trap crops	0.00 (1.04)	0.02	0.02	2,363
Burned crop residue for pest/disease control	-0.00 (-0.14)	0.06	0.06	2,363
Used traps	-0.00 (-0.31)	0.02	0.02	2,363

Note. Robust *t* statistics clustered at the clinic level in parentheses. All estimations control for lottery fixed effects and outcome value at baseline. Impact estimates are 36-month impact estimates and outcome value at baseline. Means are 36-month means. ITT = intention to treat.
*p < .10; **p < .05; ***p < .01.

Table 4.11 Impacts on Cultural Practices and Input Use: Perennial Crops (ANCOVA)

Dependent variable	ITT impact (1)	Control mean (3)	Treatment mean (4)	N (5)
Organic fertilizer used	-0.06 (-1.51)	0.68	0.63	605
Inorganic fertilizer used	-0.02 (-0.42)	0.46	0.42	605
Pesticide used	-0.08** (-2.05)	0.42	0.37	605
Used resistant variety	-0.02 (-0.39)	0.49	0.48	605
Used improved planting material	-0.04 (-0.94)	0.41	0.38	605
Used certified planting material	-0.04 (-0.71)	0.35	0.32	605
Intercrop	-0.01 (-0.14)	0.26	0.25	605
No. times crop was checked	0.18 (0.19)	11.56	11.42	605
Weeded in a timely manner	-0.04 (-1.26)	0.89	0.87	605
Removed infested or damaged material	0.04 (0.78)	0.39	0.45	605
Applied ash	0.02 (0.88)	0.08	0.11	605
Sprayed with chilies	0.00 (1.01)	0.00	0.00	605
Pruned	0.03 (0.61)	0.52	0.53	605
Changed cycle	-0.01 (-0.81)	0.04	0.04	605
Used trap crops	-0.01 (-1.35)	0.01	0.00	605
Burned crop residue for pest/disease control	-0.00 (-0.08)	0.02	0.02	605
Used traps	0.00 (0.57)	0.01	0.01	605
Greased	0.00 (0.85)	0.00	0.00	605

Note. Robust *t* statistics clustered at the clinic level in parentheses. All estimations control for lottery fixed effects and outcome value at baseline. Impact estimates are 36-month impact estimates and outcome value at baseline. Means are 36-month means. ITT = intention to treat.
*p < .10; **p < .05; ***p < .01.

Table 4.12 Impacts on Cultural Practices and Input Use - Maize Farmers (Single Difference)

Dependent variable	ITT impact (1)	Control Mean (2)	Treatment mean (3)	N (4)
Organic fertilizer used	-0.02 (-0.99)	0.45	0.42	2,161
Inorganic fertilizer used	-0.01 (-0.32)	0.82	0.82	2,161
Pesticide used	-0.02 (-0.81)	0.23	0.21	2,161
Crop rotation	0.06*** (3.03)	0.20	0.28	2,160
Used resistant variety	0.02 (1.12)	0.44	0.47	2,160
Used improved planting material	0.01 (0.27)	0.77	0.79	2,160
Used certified planting material	-0.01 (-0.39)	0.85	0.85	2,160
Removed plant residue from prior harvest	0.02 (1.32)	0.91	0.93	2,160
Planted early	0.03 (1.43)	0.71	0.74	2,160
Intercrop	-0.08*** (-3.62)	0.71	0.62	2,160
No. times crop was checked	0.54 (0.73)	12.37	13.27	2,160
Weeded in a timely manner	0.01 (0.77)	0.95	0.96	2,160
Removed volunteer crops	0.07*** (2.99)	0.58	0.65	2,160
Removed infested or damaged material	0.03 (1.18)	0.56	0.59	2,160
Mulch	-0.01 (-1.43)	0.03	0.02	2,160
Applied ash	0.02* (1.85)	0.04	0.06	2,160
Sprayed with chilies	0.01** (2.53)	0.01	0.02	2,160
Staked	0.01 (1.10)	0.04	0.05	2,160
Applied sand	0.02** (2.38)	0.04	0.06	2,160
Used trap crops	0.00 (1.13)	0.01	0.02	2,160
Burned crop residue pest/diseases control	0.01 (1.30)	0.03	0.05	2,160
Use traps	-0.00 (-0.71)	0.02	0.01	2,160
Covered with leaves	-0.00 (-0.61)	0.01	0.01	2,160

Note. Robust *t* statistics clustered at the clinic level in parentheses. All estimations control for lottery fixed effects and outcome value at baseline. Impact estimates are 36-month impact estimates and outcome value at baseline. Means are 36-month means. ITT = intention to treat.
*p < .10; **p < .05; ***p < .01.

Qualitative data indicated that farmers did not necessarily use the prescribed option, which could also partially explain the lack of impacts on intermediate outcomes associated with some input use for all crops. At the same time, farmers' inability to buy chemicals because of a shortage of funds may have influenced the significant outcomes on treatment farmers' use of cultural practices as a solution for maize. Multiple farmers indicated that money was a primary factor in deciding how to address a challenge, and that "we do short cuts because of lack of money." One farmer said, "I was told to buy the chemicals, and I used ash instead; the ash is slower in treating the banana weevil."

The last set of intermediate outcomes are the ones related to pesticide knowledge, practice, and use, given that the appropriate use of pesticides by farmers is a key component of PW-K's theory of change. Table 4.13 shows that farmers in treatment areas adopted some good practices related to pesticide use. Specifically, farmers in treatment areas are 4 percentage points more likely to check for plant health issues on a regular basis, 6 percentage points less likely to prefer chemical pest control, 3 percentage points less likely to spray pesticide in the morning, and 7 percentage points more likely to avoid chemical drift when spraying. Interestingly, treatment farmers use 0.21 fewer protective items such as gloves, masks, googles, caps, boots, and overcoats for pesticide application. While significant, this estimate is not very large given that the mean for the control group is to use 2.31 items. The results for pesticide knowledge and practice for farmers of some specific crops such as maize and beans are very similar (See Appendix 1, Tables A1.11 and A1.15). We also investigated if farmers in treatment areas reduce their number of pesticide applications and reduce the number la labor days (family and paid) used for pesticide application for all the common but did not find any impacts.[38]

[38] Estimating the effect of the program on quantity of pesticide used per unit of area is not feasible because farmers report different units (kg and liters) for different types of pesticides.

Table 4.13 Pesticide Knowledge and Practice – All Crops (ANCOVA)

Dependent variable	ITT impact (1)	Control mean (2)	Treatment mean (3)	N (4)
Pre-harvest interval is important	-0.01 (-0.38)	0.48	0.47	2,550
Check for plant health problems on a regular basis	0.04* (1.87)	0.73	0.77	2,550
Prefers chemical pest control	-0.06** (-2.56)	0.63	0.57	2,550
No. protective items for pesticide app	-0.21** (-2.15)	2.31	2.12	2,550
Spray pesticide in the morning	-0.03* (-1.85)	0.81	0.78	1,868
Spray pesticide in the evening	-0.02 (-1.26)	0.30	0.27	1,868
Avoid chemical drift when spraying	0.07*** (3.38)	0.28	0.35	1,868
Washing self after spraying	-0.00 (-0.09)	0.62	0.62	1,868
Washing equipment after spraying	-0.01 (-0.30)	0.58	0.57	1,868
Using containers only for pesticide	-0.01 (-0.31)	0.42	0.41	1,868
Chemical disposal: use it all	0.01 (0.72)	0.26	0.27	1,868
Chemical storage: shed or barn	0.01 (0.54)	0.45	0.46	1,881

Note. Robust *t* statistics clustered at the clinic level in parentheses. All estimations control for lottery fixed effects and outcome value at baseline. Impact estimates are 36-month impact estimates and outcome value at baseline. Means are 36-month means. ITT = intention to treat. *p < .10; **p < .05; ***p < .01.

Production Diversity

Table 4.14 presents results for crop diversity and production for all crops (annual, perennial, and crops planted in areas smaller than 1/32 acre) aggregated at the household level. Crop diversity and production areas were similar between the treatment and control groups. For example, the treatment group produced 8.95 crops on average, while the control group produced 9.10 crops. Both treatment and control farmers produced crops on 3.03 acres of land.

Table 4.14 Impacts on Crop Diversity: All Crops (ANCOVA)

Dependent variable	ITT impact (1)	Control mean (2)	Treatment mean (3)	N (4)
No. of all crops produced	-0.52 (-1.54)	9.10	8.95	2,550
No. of crops produced in a large area	-0.15 (-1.54)	3.94	3.86	2,550
Production area, in acres	0.03 (0.27)	3.03	3.03	2,550

Note. Robust *t* statistics clustered at the clinic level in parentheses. All estimations control for lottery fixed effects and outcome value at baseline. Impact estimates are 36-month impact estimates and outcome value at baseline. Means are 36-month means. ITT = intention to treat. *p < .10; **p < .05; ***p < .01.

We found little evidence of a change in crop diversity measures when we considered annual and perennial crops. Table 4.15 presents the crop diversity impact estimates for annual crops and perennials. Across the table, the only statistically significant estimate suggests that treatment farmers decreased the number of annual crops they produced by 0.19 crops.

Table 4.15 Impacts on Crop Diversity: Annual and Perennial Crops (ANCOVA)

Dependent variable	ITT impact (1)	Control mean (2)	Treatment mean (3)	N (4)
Annuals:				
No. of all crops produced	-0.19* (-1.84)	3.97	3.86	2,363
No. of crops produced in a large area	-0.05 (-1.21)	2.08	2.02	2,363
Production area in acres	-0.08 (-0.75)	2.70	2.58	2,363
Perennials:				
No. of all crops produced	-0.23 (-1.23)	5.69	5.48	605
No. of crops produced in a large area	-0.01 (-0.22)	1.64	1.62	605
Production area in acres	0.10 (0.67)	0.89	0.95	605

Note. Robust *t* statistics clustered at the clinic level in parentheses. All estimations control for lottery fixed effects and outcome value at baseline. Impact estimates are 36-month impact estimates. Means are 36-month means. ITT = intention to treat. *p < .10; **p < .05; ***p < .01.

We investigated if the small reduction in the number of annual crops produced was driven by specific crops. In Table 4.16, we estimate whether being in a treatment area changed the probability of producing any of the most common annual crops. The results show that treatment farmers were 1 percentage point more likely than control farmers to produce maize. There was also a small reduction in the probability of producing other common annual crops relative to the control group, but the impacts were not statistically significant. However, when we look at the combined probability of producing any of the three major annual crops except maize (i.e., beans, potatoes, and kale), we see that treatment farmers were 2 percentage points less likely to produce any of them. That means that the small reduction in the number of annual crops was not concentrated on a specific crop. In other words, the reduction in the number of small crops does not reflect a systematic substitution of some crops for maize; it simply shows that some farmers stopped producing beans, other potatoes, and other kale.

Table 4.16 Impacts on Probability of Producing Common Crops (Single Difference)

Dependent variable	ITT impact (1)	Control mean (2)	Treatment mean (3)	N (4)
Maize	0.01* (1.89)	0.18	0.20	9,797
Beans	-0.01 (-1.16)	0.17	0.16	9,797
Potato	-0.01 (-0.89)	0.06	0.06	9,797
Kale	-0.00 (-0.77)	0.04	0.04	9,797
Beans, Potato, or Kale	-0.02* (-1.80)	0.27	0.26	9769

Note. Robust *t* statistics clustered at the clinic level in parentheses. All estimations control for lottery fixed effects. Impact estimates are 36-month impact estimates. Means are 36-month means. ITT = intention to treat. *p < .10; **p < .05; ***p < .01.

Yields and Costs

Ultimately, the program aims to achieve two final impacts: improvement in production and farmer wellbeing. Data on crop production amounts and market values—along with data on input expenditures such as fertilizers, pesticides, and labor—allow us to estimate program effects on crop productivity and gross margins.

Before presenting the results, it is worth discussing some technical decisions we made to conduct the estimation. First, for the production variables, we present the estimates based on natural logarithm transformed outcome variables, in which the transformed outcome variable equals the natural log of the original outcome variable.[39] While we controlled for outliers[40] during our analysis, the natural log transformation helped further account for any outliers present with these variables. With this transformation, the interpretation of the impact of PW-

[39] Agricultural outcomes are commonly highly skewed to the right (i.e., the mean is much larger than the median) due to the presence of outliers. It is common practice to log those variables to reduce the influence of outliers.
[40] For numeric variables, we also set as missing those values above the 99th percentile and those below the 1st percentile, to further control for outliers.

K was that the outcome changed by 100%* (impact estimate), all else being equal.[41] Second, we only considered production values and input expenditures for crops cultivated in an area larger than 125 square meters (or 1/32 acre) as we did not collect data for areas below that threshold. Third, we calculated the value of yields per acre by multiplying the quantity of each crop produced per acre by the farmgate or market price of the produce at the household or village level (whichever is available in the data). Fourth, we used the value of yields, as opposed to yields, to aggregate the different crops by the type of plant life cycle (i.e., annuals or perennials). We also focused on the value of yields because around 60% of farmers in our sample use intercropping, which makes area estimates for each crop difficult to calculate. Fifth, in general, one needs to be careful when interpreting gross margin estimates because, as shown in Section 3, this variable is the sum of all production output and input values. This means that the final construction of the gross margins is very sensitive to measurement error and outliers in any of its individual components. For this reason, some authors prefer to focus on outcomes such as value of yield per unit of area when evaluating the impacts of extension programs like PW-K (Ragasa & Mazunda, 2018). Lastly, the program estimates for all variables were estimated using linear regressions.

Table 4.17 and Table 4.18 contain the production impact estimates for annual and perennial crops, respectively. For annual crops, we found an increase in the use of some cultural practices, but we did not find statistically significant estimates for production measures such as the value of yields per acre or the costs per acre (Table 4.17). We also did not observe impacts in the gross margin variable. Although the estimated program impact on the value of production is 11%, and 10% on total cost, the estimated impact on gross margins is a reduction of 54%, which is not consistent with the other two results. That is why we are cautious when interpreting the estimated impacts on gross margins.

[41] For example, suppose you have a model of the form $Log(y) = \alpha + \beta * treat + \varepsilon$, where y is production of a crop in kilograms and treat is a dummy variable equal to 1 if part of the treatment group and 0 otherwise. If the estimated value of $\hat{\beta} = 0.3$, we say that treatment increases the production of the crop in 30% (=100*$\hat{\beta}$).

Table 4.17 Impacts on Production: Annual Crops (ANCOVA)[42]

Dependent variable	ITT impact (1)	Control mean (2)	Treatment mean (3)	N (4)
Value of production per acre	0.11 (1.16)	9.24	9.39	1,494
Cost of seed planted per acre	0.09 (0.87)	7.73	7.83	1,898
Cost of inorganic fertilizer per acre	0.01 (0.04)	6.62	6.80	1,733
Cost of pesticide per acre	-0.00 (-0.02)	3.24	3.29	1,537
Cost of labor per acre	0.17 (0.89)	4.70	4.89	1,670
Total costs per acre)	0.10 (1.47)	9.09	9.21	1,901
Gross margins per acre	-0.54 (-0.95)	3.16	2.44	1,494

Note. All outcome values are in natural logs. Robust *t* statistics clustered at the clinic level in parentheses. All estimations control for lottery fixed effects and outcome value at baseline. Impact estimates are 36-month impact estimates. Means are 36-month means. ITT = intention to treat. *p < .10; **p < .05; ***p < .01.

For perennial crops, we found a statistically significant reduction in the cost of pesticide per acre (Table 4.18). For treatment farmers, the cost of pesticide per acre decreased by 97%. In Table 4.20 below, we show that this impact is not driven by one of the common perennial crops in our sample (i.e., coffee). Thus, we interpret this result as being jointly driven by the different perennials reported in the sample.

While aggregating outcome variables by type of plant life cycle (annual or perennial) allows us to characterize the agricultural outcomes of households with multiple crops, it can also prevent us from observing program impacts on some specific crops. For this reason, we also examined program impacts for some key crops, such as maize, coffee, beans, kale, and tomatoes—the most common crops in our sample and the crops most frequently brought to clinics (see Table 4.8).

[42] Note that the number of observations is different for different outcomes. This is because these variables have different numbers of missing values at baseline. Note also that if a household did not report a value at baseline but reported a value at endline, that observation is not included in the ANCOVA regression but is included in the single difference regression. The ANCOVA and single difference results are very similar (see Table A1.7 in Appendix 1).

We first looked at program impacts for the subsample of farmers who produce maize—a crop produced by 87% of farmers in our sample.[43] Table 4.19[44] presents the impact estimates for maize production and shows that treatment farmers experienced a 13% increase in the value of production per acre. The result provides evidence that PW-K leads to an improvement in productivity for farmers.

Table 4.18 Impacts on Production: Perennial Crops (ANCOVA)

Dependent variable	ITT impact (1)	Control mean (2)	Treatment mean (3)	N (4)
Value of production per acre	-0.23 (-1.01)	10.28	10.11	412
Cost of inorganic fertilizer per acre	-0.19 (-0.52)	4.71	4.37	429
Cost of pesticide per acre	-0.97** (-2.11)	4.23	3.42	424
Cost of labor per acre	0.25 (0.60)	3.79	4.18	436
Total costs per acre	0.27 (0.63)	6.37	6.52	512
Gross margins per acre	0.35 (0.60)	6.67	7.09	512

Note. All outcome values are in natural logs. Robust *t* statistics clustered at the clinic level in parentheses. All estimations control for lottery fixed effects and outcome value at baseline. Impact estimates are 36-month impact estimates. Means are 36-month means. ITT = intention to treat. *p < .10; **p < .05; ***p < .01.

In Table 4.20, we present the results for key final outcomes for other common crops in our sample. As shown, there are no impacts of PW-K on other individual crops in terms of value or yields, costs, or gross margins. The only significant impact observed for the crops considered is a 59% reduction[45] (not shown in table) in the cost of pesticide per acre for potato producers. The fact that we do not observe impacts on some of these crops is partly explained by the small sample size used for estimation. Except for beans, which are the second most commonly produced crop in our sample, with a sample size of 1,135 observations, the sample sizes for the three other crops are low. For example, program impacts for kale producers are estimated with only 60 observations. That means that we do not have enough statistical power to estimate

[43] The analysis of POMS data also shows that maize is the most common crop brought to plant clinics by farmers.

[44] Farmers changed their production patterns over time to produce more maize, which meant that the ANCOVA estimate for maize production is not appropriate. Fewer farmers produced maize at baseline, so it did not make sense to control for baseline values of maize cultural practices and input use. Instead, we report the maize estimate based on the single difference comparing endline results for treatment and control.

[45] This impact is significant at the 10% level of confidence.

program impacts even if there is an impact. There are two reasons why the number of observations by crop are small. The first reason has to do with the high level of specialization by farmers in our sample. Most farmers in our sample produce maize (88%) and beans (63%); these are followed distantly by bananas (32%), avocados (23%), and potatoes (22%). Second, we only collected detailed production data for crops produced in areas larger than 1/32 acre. This is particularly relevant for some common annual crops such as tomatoes, and perennial crops such as bananas and avocados, which are usually produced in areas smaller than the threshold.

Table 4.19 Impacts on Maize: Production (Single Difference)

Dependent variable	ITT impact	Control mean	Treatment mean	N
	(1)	(2)	(3)	(4)
Value of production per acre	0.13**	9.42	9.55	1,460
	(2.03)			
Cost of seed planted per acre	0.04	7.69	7.73	1,460
	(1.23)			
Cost of inorganic fertilizer per acre	-0.11	6.88	6.82	1,460
	(-0.62)			
Cost of pesticide per acre	-0.18	1.86	1.68	1,460
	(-1.06)			
Cost of labor per acre	-0.05	4.31	4.24	1,460
	(-0.23)			
Total costs per acre	0.03	8.94	8.97	1,460
	(0.57)			
Gross margins per acre	0.13	5.08	5.03	1.458
	(0.35)			

Note. All outcome values are in natural logs. Robust *t* statistics clustered at the clinic level in parentheses. All estimations control for lottery fixed effects. Impact estimates are 36-month impact estimates. Means are 36-month means. ITT = intention to treat. *p < .10; **p < .05; ***p < .01.

Table 4.20 Production Impacts on Other Common Crops (Single Difference)

Dependent variable	ITT impact (1)	Control mean (2)	Treatment mean (3)	N (4)
Beans				
Value of production per acre	-0.09	8.61	8.51	1,135
	(-1.42)			
Total costs per acre	-0.11	8.35	8.25	1,135
	(-1.63)			
Gross margins per acre	-0.61	1.82	1.10	1,134
	(-1.11)			
Potatoes				
Value of production per acre	-0.07	10.37	10.32	359
	(-0.62)			
Total costs per acre	-0.15	11.37	11.37	359
	(-1.18)			
Gross margins per acre	1.46	-7.75	-6.43	350
	(1.63)			
Kale				
Value of production per acre	0.76	9.41	9.02	61
	(0.99)			
Total costs per acre	-0.91	8.29	8.41	61
	(-1.12)			
Gross margins per acre	1.43	1.87	-0.18	59
	(0.55)			
Coffee				
Value of production per acre	-0.10	10.85	10.81	331
	(-0.70)			
Total costs per acre	-0.05	8.81	8.75	331
	(-0.24)			
Gross margins per acre	0.02	10.36	10.53	331
	(0.14)			

Note. All outcome values are in natural logs. Robust *t* statistics clustered at the clinic level in parentheses. All estimations control for lottery fixed effects. Impact estimates are 36-month impact estimates. Means are 36-month means. ITT = intention to treat. *p < .10; **p < .05; ***p < .01.

Potentially connected to the findings on yields and productivity, we found some evidence that treatment farmers were less likely to report changes in external production factors. Treatment farmers were less likely than control farmers to report that the timing of rainfall had changed (Table 4.21). This finding was even more pronounced for maize producers, where treatment farmers were less likely than control farmers to report that the timing of rainfall had changed, the temperature had changed, and that crop yields had decreased (Appendix 1, Table A1.26). If we assume that farmers report changes in external agricultural production factors because they feel they lack control over that process, these results suggest that PW-K may have increased farmers' sense of control over the agricultural production process.

Table 4.21 Crop Production History and Changes to External Production Factors

Dependent variable	ITT impact (1)	Control mean (2)	Treatment mean (3)	N (4)
Years HH has been farming	6.86 (1.12)	32.39	39.84	2,540
Years farming in location	8.14 (1.37)	26.49	34.91	2,544
Changed crops produced in last 5 years	0.01 (0.44)	0.23	0.23	2,421
Amount rainfall changed	-0.00 (-0.10)	0.87	0.87	2,421
Timing rainfall changed	-0.03*** (-2.96)	0.95	0.92	2,421
Temperature changed	-0.02 (-1.28)	0.86	0.85	2,421
Number of insects increased	-0.02 (-0.76)	0.56	0.54	2,421
Number of diseases increased	-0.02 (-0.93)	0.55	0.52	2,421
Crop yields decreased	-0.03 (-1.44)	0.71	0.68	2,421
Pest information improved	0.04** (2.42)	0.30	0.34	2,421

Note. Robust t statistics clustered at the clinic level in parentheses. All estimations control for lottery fixed effects. Impact estimates are 36-month impact estimates. Means are 36-month means. ITT = Intention to Treat. *p < .10; **p < .05; ***p < .01.

Wellbeing

To examine the impact of PW-K, we focused on impacts on food security, based on actual reports of specific food items consumed. In Tables 4.22 and 4.23, we report the impact estimates for specific food items consumed to get a sense of whether increased maize production has led to greater dietary diversity. We used a household dietary diversity measure developed by the Food and Nutrition Technical Assistance Project (FANTA), which calculates dietary diversity scores by summing the number of food groups consumed by anyone in the household during a reference period (our questionnaire used the last 7 days).

We do not find any overall impacts on the household dietary diversity score for all households or for maize-producing households. However, we did find some evidence of positive increases in consumption for some food groups: when looking at all farmers, the treatment group was 2 percentage points more likely to consume fruit, 3 percentage points more likely to consume red meat and poultry, and 1 percentage point more likely to consume condiments than the treatment group (Table 4.22). The results were similar when we considered only households that produce maize (Table 4.23). The lack of impact on the overall dietary diversity score is not surprising, given that the number of total annual crops was slightly reduced and the observed increase in production occurred in maize—a crop that most households are already consuming.

Table 4.22 Household Dietary Diversity Score (Single Difference)

Dependent variable	ITT impact (1)	Control mean (2)	Treatment mean (3)	N (4)
Grains	-0.00 (-1.07)	0.99	0.99	2,550
Tubers	0.02 (1.63)	0.93	0.94	2,550
Vegetables	-0.00 (-0.12)	0.99	0.99	2,550
Fruits	0.02** (2.05)	0.86	0.89	2,550
Red meat and poultry	0.03* (1.67)	0.62	0.64	2,550
Eggs	0.00 (0.20)	0.70	0.70	2,550
Fish	0.02 (1.14)	0.24	0.25	2,550
Legumes	0.01 (0.71)	0.94	0.95	2,550
Milk	0.01 (0.67)	0.89	0.89	2,550
Fats	-0.00 (-0.06)	0.89	0.88	2,550
Sweets	0.00 (0.02)	0.92	0.92	2,550
Condiments	0.01* (1.83)	0.96	0.98	2,550
HH (household) dietary diversity score	0.08 (1.36)	9.98	10.05	2,538

Note. Robust *t* statistics clustered at the clinic level in parentheses. All estimations control for lottery fixed effects. Impact estimates are 36-month impact estimates. Means are 36-month means. ITT = intention to treat. *p < .10; **p < .05; ***p < .01.

Table 4.23 Household Dietary Diversity Score: Maize Producers Only (Single Difference)

Dependent variable	ITT impact (1)	Control mean (2)	Treatment mean (3)	N (4)
Grains	-0.00 (-0.41)	0.99	0.98	1,877
Tubers	0.02* (1.90)	0.93	0.96	1,877
Vegetables	-0.00 (-0.40)	0.99	0.99	1,877
Fruits	0.03** (2.06)	0.87	0.90	1,877
Red meat and poultry	0.04* (1.93)	0.60	0.65	1,877
Eggs	0.00 (0.01)	0.72	0.72	1,877
Fish	0.02 (0.88)	0.23	0.25	1,877
Legumes	-0.00 (-0.28)	0.95	0.95	1,877
Milk	0.01 (1.10)	0.88	0.89	1,877
Fats	-0.01 (-0.44)	0.90	0.90	1,877
Sweets	-0.00 (-0.25)	0.92	0.92	1,877
Condiments	0.00 (0.48)	0.97	0.97	1,877
HH dietary diversity score	0.08 (1.10)	9.99	10.08	1,872

Note. Robust *t* statistics clustered at the clinic level in parentheses. All estimations control for lottery fixed effects. Impact estimates are 36-month impact estimates. Means are 36-month means. ITT = intention to treat. *p < .10; **p < .05; ***p < .01.

D. Are the additional costs of PW-K justified given the benefits the program provides?

When assessing the impact of a program, a key question is whether the monetary gains created by the intervention outweigh the program running costs. To compare projects with different time lengths, we need to aggregate the project's benefits and costs over time. However, aggregating the annual flows of benefits and costs over time needs to account for the fact that a dollar today is worth more than a dollar tomorrow or in 10 years' time. To aggregate monetary values for different years, we need to express all past and future flows of benefits and costs in present value terms, using a discount rate. A common discount rate is the market interest rate.

We used two common measures to assess whether the benefits of PW-K justify the costs. The first measure was the benefit-cost ratio, which is given by the share of the present value of benefits to the present value of costs:

$$BC = B/C = \frac{\sum_{j=0}^{T} B_j/(1+i)^j}{\sum_{j=0}^{T} C_j/(1+i)^j}$$

Where B is the present discounted value (PDV) of the program benefits from the initial year of the program (i.e., when j = 0) up to a future year T. The PDV of the benefits is calculated by adding the yearly benefits of the program after discounting each year's flow using the interest rate (i). The PDV of the costs is calculated in a similar way. According to the benefit-cost ratio, an investment is profitable if the ratio is greater than 1—in other words, if benefits are larger than costs for the duration of the project.

The second measure we used to assess the program's profitability is the internal rate of return (IRR). This is defined as the discount rate that yields the PDV of the net benefits (i.e., benefits minus costs) equal to zero. That is:

$$0 = \sum_{j=0}^{T} (B_j - C_j)/(1+IRR)^j$$

According to the IRR criterion, an investment is profitable if the computed IRR is greater than the market interest rate of return.

To calculate these two measures of program profitability, we first needed to calculate the program costs and benefits for a determined period.[46] [47] In order to calculate PW-K program costs, we used the ingredients approach. This approach is a systematic, well-tested procedure for identifying all comprehensive costs for implementing program services, including costs that are routinely not adequately identified in budget or expenditure data, such as contributed (in-kind) resources, opportunity costs, or costs that are shared between the program and other operational activities.

[46] For the analysis, we define time 0 as 2012, the initial year of Plantwise in Kenya; and time T, the final year of the analysis, as 2024. We chose 2024 as the final year in order to have a 10-year time period from the moment we assume PW started producing benefits for farmers.

[47] While the choice of the final year in the estimation is arbitrary, the profitability estimates are highly insensitive to changes in the final year. This is because we are discounting benefits and costs over time to express all magnitudes in terms of 2012 values. Thus, the contribution of a given year to the estimation of the profitability measures decreases as time goes by. For example, the contribution of the magnitudes in 2024 to the analysis are lower than the contribution of the magnitudes in 2023 and so on. This means that the results do not vary importantly if we assume a different final year for the estimation close to 2024.

The costs associated with PW-K fall into three main categories: (1) CABI and national coordination and advocacy, (2) plant clinic operations, and (3) the Knowledge Bank and POMS operations (see Table 4.24). For each of these sets of activities, the additional costs of PW-K (beyond the normal operating costs of the agricultural extension system) include the costs of investing in each of these activities—both to initiate the activities and to maintain them—as well as the opportunity costs of government employees' time.

Table 4.24 PWK Cost Analysis 2015–2017

Cost analysis 2015–2017				
Coordination, advocacy, and M&E	**Source**	**2015**	**2016**	**2017**
National-level support activities	CABI	736	4,556	233
Advocacy activities	CABI	17,716	2,954	2,040
Advocacy activities	GoK	5,391	1,662	41
CABI coordination	CABI	105,526	113,935	78,467
Salary of key Plantwise staff	GoK	4,053	4,715	5,147
Plantwise key staff – additional funding	CABI	20,315	7,543	4,831
M&E	CABI	5,043	8,871	9,358
Subtotal (GBP)		**158,779**	**144,235**	**100,118**
Plant clinics	**Source**	**2015**	**2016**	**2017**
Clinic operations	CABI	21,694	21,268	21,681
Local coordination	CABI	8,310	0	7,049
PD trainings	CABI	66,289	44,785	6,996
Salary of PDs	GoK	218,570	217,222	235,210
Salary of clinic coordinators	GoK	9,044	5,810	5,810
Clinic coordinator and PD additional funding	CABI	31,757	51,762	45,249
Subtotal (GBP)		**355,664**	**340,848**	**321,996**
POMS and Knowledge Bank	**Source**	**2015**	**2016**	**2017**
Salary of data manager	GoK	9,849	12,479	3,242
Data manager – additional funding	CABI	196	398	1,025
Data entry, validation, and harmonization	CABI	7,887	6,906	4,995
Materials (tablets/fulcrum)	CABI	5,628	1,238	9,100
Curriculum development	CABI	3,435	0	260
Diagnosis	CABI	0	16	268
Knowledge Bank support	CABI	13,534	6,639	4,925
Knowledge Bank costs*	CABI	96,338	82,851	85,741
Subtotal (GBP)		**136,867**	**110,526**	**109,555**

* We assume that the proportion of the global Knowledge Bank costs that are assigned to Kenya are 10%, 8.6%, and 8.9% for each year between 2015 to 2017, respectively.

Several costs are associated with getting PW started in Kenya and maintaining its organization. First, we included CABI coordination and national coordination costs associated with organizing and participating in the national forums and steering committee meetings. Government costs for national coordination are the estimated time costs for government employees who attended the meetings. These are based on salary estimates published by the Kenyan government in 2016, which are still in effect. These costs run on assumptions about the number of days for each meeting and the job group levels of attendees. Additionally, this category includes costs for other advocacy activities, including marketing PW-K. These costs vary by year, depending on the extent of activities. To ensure effective operation of PW-K, the MoA employs one person whose time is completely devoted to PW-K (working closely with CABI), and a manager who dedicates a significant portion of their time to PW-K activities. Both positions receive salary top-ups from CABI to cover additional costs of activities associated with PW-K. Lastly, we include the costs of M&E activities undertaken by CABI researchers, including a pesticide study and gender study.

The next category of costs is associated with plant clinic operations. Costs within this category include those related to initiating plant clinics: the costs of training plant doctors, the costs of materials required to set up and operate the clinics, and the costs of local coordination. As plant doctors and clinic coordinators are employed as extension officers with the MoA, they are provided their regular salaries through the Kenyan government, according to the salary rates mentioned above. They, too, receive additional funding through CABI to support their work with PW-K: clinic coordinators receive monthly salary top-ups, while plant doctors receive money for airtime and travel costs for relevant PW-K activities.

Once the plant clinics are established and operational, data from the clinics are collected, validated, and organized within POMS. Accordingly, costs in this final category include those associated with updating and maintaining the data management system, and the costs of equipment necessary for data collection (such as tablets and fulcrum). A data manager, employed by the Kenyan government, provides oversight for all data entry, validation, and harmonization activities. This data manager receives salary top-ups from CABI to support the work they do for PW-K. They also assist with curriculum development for trainings and diagnoses of plant diseases or pests. Lastly, this category encompasses funding to support Knowledge Bank activities in Kenya, both directly and indirectly, as a percentage of worldwide Knowledge Bank support expenditures.

Costs for all three categories were totaled by year and funder, as seen in Table 4.25. All cost information used in this analysis was provided by CABI for the period 2012 to 2017.

Table 4.25 Total PWK Costs by Funder, 2015–2017

Total costs by year (GBP)			
	2015	**2016**	**2017**
Total	651,310	595,609	531,669
Total CABI	404,404	353,721	282,218
Total GoK	246,906	241,888	249,451

Program benefits were calculated from the estimated results of the impact assessment (Figure 4.4). The calculation of program benefits focused exclusively on maize outcomes, as this is the crop for which the evaluation found an economically and statistically significant impact on the value of production. While it is possible that PW-K is generating positive impacts for other crops, the results for other crops were not statistically significant. We estimated program benefits by multiplying the 13% increase in the annual value of production for the average maize farmer by the value of maize production per acre (KSH 16,200). We then multiplied that by the number of acres cultivated in maize per farmer (1.33 acres) and the number of farmers in a plant clinic catchment area who cultivate maize (636), and then multiplied by 122 (the total number of plant clinics in 2017).

Figure 4.4 Benefit Calculation

Benefits = *(A x B)* *x* *(C x D)* *x* *E*

A = % increase in yearly gross margins for maize farmers due to PW **= 13%**

B= Value of maize production per acre = 540kg x KES 30/kg = KES 16,200 **= GBP 113.4**

C = No. of acres cultivated in maize per farmer **= 1.33 acres / farmer**

D = No. of farmers in PC catchment area who cultivate maize **= 636 farmers**

E = No. of plant clinics **= 122 PCs**

Total program costs in 2017 were estimated to be GBP 531,669, and total program benefits in 2017 were estimated to be GBP 1,521,335. This gives a benefit-cost ratio for 2017 of 1,521,335/491,270, or approximately 2.8:1, showing that the benefits considerably outweighed the costs of running the program in 2017. Assuming that the costs and benefits remain stable after 2017, the benefit-cost ratio for the 2012–2024 period is 2.1:1—just slightly lower than the 2017 benefit-cost ratio. As in most development projects, this is because costs during the first years of the program were higher than the benefits.

We also calculated the associated IRR of PW-K to be 54%, using the methodology introduced above. As discussed, the higher a project's IRR, the more desirable it is to undertake a project. For the evaluation of PW-K, the IRR was estimated using the following assumptions: (1) The number of plant clinics will remain stable for the period 2018 to 2024; (2) it takes 2 years for a plant clinic to start generating the observed monetary benefits we estimated in 2017; (3) there were no monetary benefits in 2012 and 2013; and (4) program benefits and costs will remain stable in real terms for the period 2018–2024.

At present, CABI funds most of the investment into PW-K through direct payments and the time of their staff. The opportunity costs of MoA staff time are covered by the Kenyan government. However, in the future many of these costs are unlikely to exist as systems become more developed; if the program is sustainable and the government absorbs the running costs of PW-K, many costs would be unnecessary. We therefore conducted an additional benefit-cost analysis, factoring in only those costs and activities that would be undertaken by the government should CABI transition out of ownership. If costs related to CABI coordination, national coordination, and advocacy are excluded, the total annual costs would become approximately 65% of current program costs and the benefit-cost ratio for 2017 would increase to 3.5:1.

The estimated measures used to assess the profitability of PW-K show that the program provides good value for money. First, compared to cost-benefit analyses conducted on other agricultural extension programs, the estimated IRR for PW-K is above average. A systematic review by the International Food Policy Research Institute in 2000 (Alston et al., 2000) found the median IRR for similar research and extension programs to be 37%, which may be an overestimate due to publication bias and the quasi-experimental nature of most studies. A 2015 systematic review of the effects of training, innovation, and new technology on African smallholder farmers by the Campbell Collaboration (Stewart et al., 2015) shows that although there are some positive indications that training interventions might have beneficial effects on farming households' income, these findings are not statistically significant, which means that benefits of programs in Africa similar to PW-K are not different from zero.

Second, program benefits are likely to be underestimated, and costs are likely to be overestimated, as they include CABI program-level inputs for running a program that includes research elements. The analytic method used for the evaluation was also not able to measure the explicit effects of PW-K on yield and costs of production for crops grown on small plots of land, such as the tomatoes, kale, and horticultural crops commonly brought to clinics. In addition, plant health systems changes are expected to deliver other monetary benefits in Kenya, such as being able to identify new pests at the national level. Impact assessments are not able to capture these impacts because both treatment and control farmers are positively affected by such improvements, and program benefits are therefore likely to be underestimated. This is potentially the case for the MoA's response to the fall armyworm outbreak, which used support from PW. This may have positively affected both treatment and control areas alike.

Finally, as is common in other development programs, benefits may increase relative to costs over time, as knowledge learned by farmers and other stakeholders is reused without further need for direct advice on recurrent problems.

5. Conclusions, Implications for PW-K, and Next Steps

In this section we discuss the key findings of the evaluation and provide some recommendations based on the data we collected over the four years of the evaluation. The conclusions and recommendations are presented using the four research questions that motivated the evaluation.

A. Plant Health System Change

PW-K is improving institutional coordination in the plant health system, generating more knowledge, and subsequently improving the likelihood of identifying outbreaks and responding to them in a timely manner. Stakeholders perceived that PW-K has altered how farmers interact with MoA entities at the local level. Plantwise provides critical agricultural extension services that the county otherwise could not have because of weaknesses at the county level that respondents attributed to devolution, lack of funding, and a shortage of staff. The key findings from key informant interviews were as follows:

The interaction with farmers through plant clinics was widely viewed as helpful in addressing farmers' needs concerning plant health. This is, to a large extent, due to the training of plant doctors and improved access to information. When farmers were asked where they would go for information and advice if plant clinics did not exist, very few said they would consult the MoA.

Plant clinics were mentioned as the primary way of identifying pests and diseases. Most respondents said there was no established government system for detecting or reporting new pest invasions, and that Plantwise reporting systems filled a significant gap in capturing data and reporting on agricultural information related to pests and diseases. Overall, stakeholders said they valued Plantwise data for their ability to track disease outbreaks. Nevertheless, the interview results indicated that there is still room for local stakeholders to increase their use of PW-K systems to take full advantage of the data collected.

Plant clinic activities have become a regular part of government activities. PW-K is a widely known program among county-level MoA officers. It provides concrete tasks and expectations for stakeholders and therefore is a primary focus of employees and users in the system. Plantwise activities are among the few consistent, ongoing efforts for MoA officers at the county level. Extension officers explained that they did not have regular duties aside from those for externally funded projects.

Officers in our sample cited devolution as a variable that adversely affected their work, the overall plant health system in Kenya, and Plantwise in particular. Counties, as new institutions with limited experience in providing services to stakeholders, may not yet have the organizational capacity to undertake the work they have inherited from the national government. Another problem stemming from Kenya's devolved government is the lack of consistency in stakeholders' messages concerning best agricultural practices. County officers also said that they rarely were allocated the funding they requested from the national level MoA and that local governments often prioritize sectors with higher visibility projects than agriculture. Nevertheless, despite the challenges, county and subcounty respondents found a synergy between their new responsibilities after devolution and the support offered by Plantwise.

In summary, we found that the information generated from the plant clinics through the POMS, combined with improved institutional coordination, has the potential to continue to shift the plant health system in a positive direction. The direct contact with farmers and the general shifts in the plant health system are consistent with a strengthening of the institutions that manage plant health, as well as a reduction in damage from pests and diseases in Kenya.

Recommendations: Plant Health System Change

1. **Buy-in and uptake.** Sustainability of Plantwise lies, in large part, in the ability of PW-K to secure buy-in from and uptake by mid-level county officers, who currently do not have much interaction with the Plantwise activities on the ground, nor do they work regularly with the data generated by the plant clinics into the POMS. Working closely with mid-level officers to show a strong connection between plant clinics and high priorities such as food security and production has the potential to increase investment in Plantwise activities and the likelihood that they will continue on government funding alone.
2. **Partnership with MoA.** Endline data reinforced the need to ensure that the MoA contributes activities where necessary for a complimentary approach to implementation. PW-K was not as widespread or as productive as it could be given the challenges regarding inputs the government has been expected to provide, such as internet bundles in offices, funding for employee transport to the field, and staff accountability for their work. In addition to planning for larger goals—such as building an early warning system—Plantwise should utilize the national-level representatives to ensure that assumed elements of the system are taking place.
3. **Information systems.** Plantwise should refine systems to provide timely information from the plant clinics to key actors in the plant health system to encourage rapid action and

response. Overall, implementers do not regularly communicate with decision makers enough to ensure that ground-level challenges are quickly and efficiently communicated and addressed. The current system of data collection, entry, validation, and reporting is neither regular nor sufficiently institutionalized to facilitate early warning or even a rapid response to identified disease. Since the use of e-plant clinics expanded, data collection has improved, though some new problems emerged as a result, such as connectivity issues and upload inconsistency. We therefore recommend the following actions to help refine systems of information provision:

- Encourage the use of email and formalize the use of telegrams for regular communications on data challenges, data inquiries, and diagnosis questions.
- Formalize communications through monthly calls at the district level for updates on specific items as necessary. Districts can then be required to share their challenges, tasks, action items, and next steps to the assigned MoA representative at the national level.
- Promote data-based action on the part of decision makers through protocols that outline a quick, coordinated response to potential outbreaks and chronic issues as soon as they have been identified in the field.

4. **Capacity building.** PW-K should focus on building capacity for county officers who work on PW-K in financial and logistics management—key areas of weakness highlighted by respondents in the qualitative data. Plantwise also should continue to help build systems of coordination across county institutions and with national bodies.

B. Evaluating the Implementation of PW-K

The process through which PW-K is implemented is innovative and comprehensive. In principle, it should improve knowledge at multiple levels through accessible diagnosis for farmers, improved training for extension officers, and data collection to help understand where diagnosis could be improved in the short term and where the system should address problems in the long term.

Overall, the qualitative results from the process evaluation endline data collection suggest that PW-K is one of the central components of the county-level agriculture activities implemented through the MoA. Farmers and plant doctors appreciate the knowledge that Plantwise has made available to aid pest and disease identification and diagnosis, and officers from the county level to the national level said that Plantwise data are among the most accurate and reliable.

PW-K trainings produce a large effect on knowledge generation.

The plant doctor training has had a large and statistically significant effect on plant health knowledge and was highly regarded by every stakeholder we interviewed. As the quantitative assessment results showed, extension agents who received training learned essential information about maintaining plant health that others are not aware of (e.g., “We learned that we have an active ingredient in chemicals”).

There is potential to expand the utilization and utility of POMS data.

We found that county-level officers recognized the value of POMS data for tracking disease outbreaks, and higher-level external stakeholders (e.g., national-level officers) only occasionally inquired about the data from the system. Though awareness of the existence of the POMS seems to have increased, we found that that POMS data are still not widely used. Although higher level external stakeholders only occasionally asked to receive information from the system, some were not well versed in navigating the system. Data also seemed to be used in an ad hoc manner; there was no systematic, consistent approach across counties to employing the data. Respondents attributed low use of POMS data to challenges with simple solutions, including not having access to passwords and lacking a basic understanding of how to use the data.

Collecting and using actionable data is a key aspect of PW-K. The data management system has the potential to facilitate a broader range of analyses than those currently undertaken. Officers suggested that regular and continuous validation could ensure that the POMS serves as an early warning system for plant health problems, and that plant doctors could participate in the validation process to learn from mistakes in diagnosis and about costs of recommendations.

While farmer awareness of plant clinics can be improved, clinic usage is expected.

In the 12-month period between July 2016 and July 2017, 53% of treatment farmers were aware of plant clinics. Of those, 34% of treatment group participants had attended a plant clinic at least once in the 12 months before the survey. While there is room for increasing farmer awareness of plant clinics, the plant clinic attendance levels are not unusual for a public, demand-driven program like PW, in which only a fraction of the potential beneficiaries who are offered the intervention are expected to directly participate in program activities. In fact, the most frequently cited reason for not attending plant clinics was that farmers did not feel a need to do so. Getting information from other farmers is a key reason that program implementation assessment needs to extend beyond measuring plant clinic attendance: There are other ways in which PW-K potentially can benefit farmers who live in a plant clinic catchment area.

Recommendations: Implementation of PW-K

Though activities implemented through PW-K are well regarded, useful, and productive, they are hindered by a lack of complementary facilitation by the MoA, particularly in the form of funding for basic operations as discussed in section A. In addition, the activities lack consistent follow-up that could increase the likelihood of Plantwise achieving larger impacts. For example, the POMS data do not appear to be utilized often enough to consistently facilitate a timely response to plant health outbreaks. The potential of the POMS is further limited by access challenges and infrequency of validation. The lack of continuous validation may decrease the potential of the POMS as an early warning system for plant health problems, though networks established through PW-K have enhanced the likelihood of early identification. Thus, we recommend the following:

5. **Utilize the National Forum.** The National Forum presents an opportunity for Plantwise to emphasize that the success of Plantwise activities depend on the proper functioning of the complementary activities for which the MoA is responsible, including reliable funding, diagnosis activities through KEPHIS and KALRO, and enforcement of performance standards for district- and county-level employees. National Forum representatives are also in a place to communicate to their MoA peers that Plantwise is concerned with to national priorities such as food security. National-level support is necessary to enable the funding of district- and county-level activities that are crucial enough to determine the success of Plantwise.
6. **Continue POMS data collection and validation but implement strategies to reduce cost.** Stakeholders indicated that POMS data collection and validation is extremely useful but that the current approach is too costly. To increase the usefulness of the POMS, we recommend that CABI take the following actions:
   - Conduct regular trainings at the local level, grant system access to more county-level MoA officers so they can use (not necessarily edit) the data, and incorporate simple validation checks that users are continuously prompted to perform. For example, involving plant doctors in data validation efforts could help them learn from mistakes and therefore continuously improve without the need to complete formal refresher trainings.
   - Ensure that the information collected through the POMS is regularly transmitted to key decision makers at the district and national levels. A broader range of analyses could help answer key questions; for example, a representative from the seed industry noted that the industry relies on access to accurate data and that POMS data could assist their efforts.

- Explore the possibility of budgeting for data validation at the county level to achieve greater cost efficiency (in terms of both time and money) at the national level. To complement this approach, a national-level representative or expert could be tasked with supervising data validation.
- Other structures already inherent in the Plantwise process have the potential to improve the process from the ground up. For example, it may be useful to involve plant doctors in the validation process to ensure that it occurs more frequently and that the plant doctors learn from their own experiences and those of others. Similarly, Plantwise could incorporate a follow-up element into the reporting requirements to encourage county officers to utilize the information collected for productive purposes beyond the reporting period. Finally, encouraging integration of Plantwise activities into other agriculture activities, and emphasizing ownership at the district level, may help ensure that the new capacities of Plantwise continue beyond the project period.

7. **Reconsider the selection of plant doctors based on longevity and, potentially, on recommendation quality.** Because plant doctor knowledge is linked to both training and practice, choosing individuals nearing retirement may not be the most cost-effective approach. Accordingly, we recommend that PW-K train individuals who are a certain number of years (e.g., more than 5 year) from retirement.

8. **Institute continuous plant doctor training, based on POMS analysis, and improved information provision.** The plant doctor assessment showed that long-term plant doctors perform better than other agricultural extension officers. However, plant doctors continually face unfamiliar problems and need regular access to updated information. We recommend that PW-K undertake the following measures:
   - Provide regular trainings—perhaps in the form of a train-the-trainers model—to help mitigate deterioration in plant doctors' knowledge over time and to keep them informed of emerging issues.
   - Use POMS data to identify weaknesses among plant doctors and to determine actions that help them overcome shortcomings.
   - Given that farmers are unlikely to implement the recommended remedy if they are unable to afford it, we suggest that PW-K train plant doctors to understand the costs of the different options they are proposing as solutions to the farmers.

C. Farm-Level Impact

Overall, the quantitative results of the impact evaluation at the farm level suggest that PW-K contributed to improvements in intermediate outcomes and some final outcomes. We found improvements in the use of cultural practices and inputs for farmers in areas with access to plant clinics. The evidence suggests that farmers in treated areas are applying some positive practices as a result of the program. We also learned that treatment farmers are less likely to report changes in external production factors such as reporting that the timing of rainfall has changed, that the temperature has changed, or that crop yields have decreased over time. These results can be interpreted as PW-K increasing farmers' sense of control over the agricultural production process as a result of the services they receive from the plant clinics.

We also investigated the results of the program on agricultural production for the most relevant crops in the sample. One key finding of the evaluation is that maize farmers in the treatment group experience a large and statistically significant increase in the value of maize per acre over farmers in the control group. Of the maize farmers who demonstrated improved cultural practices and inputs, those in the treatment group experienced a 13 percent increase in the value of maize per acre over farmers in the control group. This result is highly relevant given that maize is the most commonly produced crop by farmers in our sample and Kenya more broadly. The evaluation results provide strong evidence that Plantwise improved the conditions of farmers who live in plant clinic catchment areas in terms of higher crop yields and safer and better production practices.

Recommendations: Farm-level Impact

While these positive finding are promising, the qualitative results indicate that there are still some areas related to plant clinic operations that can be strengthened for developing the clinics as legitimate interfaces among the different organizations that are part of the plant health system and the farming community. To this end, we offer the following recommendations:

9. **Help farmers make better use of plant clinics by making activities more farmer-centric.** PW-K should implement a more systematic approach to marketing plant clinics to local farmers and ensure that marketing efforts target farmers' specific needs. Interviews indicated that it may be helpful to market the clinics as resources for general information provision, because many farmers find it useful to have a general conversation about their crops. We recommend the following marketing and communications strategies to increase farmers' use of clinics:

- PW-K should utilize radio announcements to distribute information to farmers, given that radio is a common source of information for farmers. This marketing strategy could incorporate the use of mobile phone technology to expand the reach of the program.
- PW-K should aim to print information for farmers, who still value and have more consistent access to paper materials. One farmer remarked, “I wish there is a calendar explaining when the diseases come in. Right now, we have no calendar showing that, so we can’t prevent diseases.”
- Despite the central location of the clinics, farmers indicated that they still value visits from extension agents. Extension agents should conduct farm follow-up visits on off-clinic days for farmers whose problems need additional attention.

  Farmers indicated that cost is the number-one reason they would not implement a recommendation. Though plant doctors said they make efforts to prescribe cultural practices or chemical solutions at the lowest possible cost, PW-K should encourage plant doctors to investigate whether any currently subsidized products might be effective solutions. Moreover, it would be useful for Plantwise to know the costs of the different recommendations they are giving to farmers.

D. Cost Analysis and Sustainability

As part of the impact evaluation of PW-K, we assessed how the benefits compare to the costs of implementing the program. We calculated two different measures to assess PW-K’s profitability, the benefit-cost ratio and its internal rate of return. The estimated measures provided to assess the profitability of PW-K show that the program provides good value for money. First, the estimated IRR for PW-K is above the returns of similar research and extension programs. Second, the benefit-cost ratio of the program is significantly greater than one, even in the most conservative case where initial set-up costs are fully included. The estimated profitability measures show that PW-K provides a good value for money, in clear contrast with recent evaluations of similar extension programs, which have not exhibited impacts on farmer-level outcomes.

Recommendations

10. **Develop a sustainability plan and begin to put mechanisms in place to ensure uptake beyond the program period.** MoA officers at all levels have not conceptualized the program as one that should become MoA owned and, therefore, have no concrete plans for

integrating the activities into their work or sustaining them with county funds. PW-K could take the following actions to encourage local uptake and sustainability:

- Multiple stakeholders recommended trying to gain buy-in from a politically influential individual who understands agriculture and can convince higher level ministers of the importance of funding the program. One stakeholder recommended explaining the importance of agriculture by connecting it with food security—a bigger issue that the non-agriculture ministry stakeholders are more likely to understand. Another stakeholder recommended communicating the value of Plantwise in understanding and predicting disease outbreaks in order to stress the importance of continuing Plantwise activities.
- Given CABI's efforts to involve national-level MoA officers and lower level extension staff, filling the missing component of mid-level involvement may support integration of the program into extension operations at multiple levels. This strategy could encourage program ownership among stakeholders at higher levels than plant doctors, thus increasing the possibility of increased allocation of government funding for the program activities.
- Present PW-K activities as a general approach to extension, rather than as program-specific activities, to raise awareness that the county offices need to provide funding specifically allocated to the program—that this funding will not come from a donor (e.g., use the term "plant clinics," rather than "Plantwise program," when referring to an approach to extension, given that there may be an inability to separate the clinics from CABI funding. Uptake as a general, regular activity should be formalized and encouraged.
- During National Forum events, ask specific questions about data coordination to gain a better understanding of how PW-K data can be institutionalized and enhanced.
- Refine the core understanding of how reporting and data are expected to be used and coordinated at the national and county levels.
- Sell the idea of PW-K data as useful for broader issues that are pertinent to non-agriculture officers, such as food security.

References

Alston, J.M., Chan-Kang, C., Mara, M.C., Pardey, P.G., & Wyatt, T.J. (2000). A meta-analysis of rates of return to agricultural R&D: Ex pede Herculem? Research Report 113, International Food Policy Research Institute, Washington, DC, USA

American Institutes for Research. (2015). *Evaluation of Plantwise–Kenya: Baseline report*. Washington, DC: Author.

Danielsen, S., Boa, E., Mafabi, M., Mutebi, E., Reeder, R., Kabeere, F., & Karyeija, R. (2013). Using plant clinic registers to assess the quality of diagnoses and advice given to farmers: A case study from Uganda. *The Journal of Agricultural Education and Extension, 19*(2), 183–201.

Denzin, N. K. (1978). *Sociological methods.* New York, NY: McGraw-Hill.

Duflo, E., Glennerster, R., & Kremer, M. (2008). Using Randomization in Development Economics Research: A Toolkit." T. Schultz and John Strauss, eds., Handbook of Development Economics. Vol. 4. Amsterdam and New York: North Holland.

Guba, E. G., & Lincoln, Y. S. (2005). Paradigmatic controversies, contradictions, and emerging confluences. In N. K. Denzin & Y. S. Lincoln (Eds.), *The Sage handbook of qualitative research* (3rd ed.) Thousand Oaks, CA: Sage.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry.* Newbury Park, CA: Sage.

McKenzie, D. (2012). Beyond baseline and follow-up: The case for more T in experiments. *Journal of Development Economics*, vol. 99(2), 210-221.

Ragasa, C., & Mazunda, J. (2018). The impact of agricultural extension services in the context of a heavily subsidized input system: The case of Malawi. *World Development*, 105, issue C, 25-47.

Stewart, R., Langer, L., Rebelo Da Silva, N., Muchiri, E., Zaranyika, H., Erasmus, Y., Randall, N., Rafferty, S., Korth, M., Madinga, N., & de Wet, T. (2015). The Effects of Training, Innovation and New Technology on African Smallholder Farmers' Economic Outcomes and Food Security: A Systematic Review. Campbell Systematic Reviews 2015:16

Appendix 1. Additional Impact Estimates

Table A1.1 Impacts on Cultural Practices: All Crops (Single Difference)

Dependent variable	ITT impact (1)	LATE impact (2)	Control mean (3)	Treatment mean (4)	N (5)
Organic fertilizer used	-0.02 (-1.04)	-0.29 (-1.02)	0.58	0.55	2,526
Inorganic fertilizer used	-0.01 (-0.67)	-0.14 (-0.67)	0.82	0.82	2,526
Pesticide used	-0.01 (-0.64)	-0.18 (-0.65)	0.49	0.49	2,526

Note. Robust *t* statistics clustered at the clinic level in parentheses. All estimations control for lottery fixed effects. Impact estimates are 36-month impact estimates. Means are 36-month means. ITT = Intention to Treat. LATE=Local Average Treatment Effect, *p < .10; **p < .05; ***p < .01

Table A1.2 Impacts on Cultural Practices: Annual Crops (Single Difference)

Dependent variable	ITT impact (1)	LATE impact (2)	Control mean (3)	Treatment mean (4)	N (5)
Organic fertilizer used	-0.03* (-1.72)	-0.40 (-1.63)	0.47	0.43	2,424
Inorganic fertilizer used	-0.01 (-0.47)	-0.10 (-0.48)	0.82	0.82	2,424
Pesticide used	-0.01 (-0.54)	-0.12 (-0.54)	0.42	0.41	2,424
Crop rotation	0.04* (1.96)	0.48* (1.88)	0.27	0.31	2,424
Resistant variety	0.02 (1.22)	0.25 (1.20)	0.45	0.48	2,424
Used improved planting material	0.02 (0.83)	0.20 (0.83)	0.76	0.79	2,424
Used certified planting material	-0.02 (-0.86)	-0.21 (-0.88)	0.84	0.83	2,424
Remove plant residue from prior harvest	0.00 (0.30)	0.04 (0.30)	0.94	0.94	2,424
Plant early	0.02 (1.20)	0.25 (1.21)	0.74	0.76	2,424
Intercrop	-0.08*** (-2.97)	-0.90** (-2.56)	0.66	0.58	2,424
Weed in a timely manner	-0.00 (-0.48)	-0.04 (-0.49)	0.98	0.97	2,424
Remove volunteer crops	0.05** (2.16)	0.55** (2.05)	0.66	0.70	2,424
Remove infested or damaged material	0.04* (1.85)	0.45* (1.84)	0.61	0.65	2,424
Mulch	-0.01 (-1.12)	-0.14 (-1.14)	0.07	0.06	2,424
Apply ash	0.03** (2.52)	0.32** (2.41)	0.05	0.08	2,424
Spray with chilies	0.01*** (2.82)	0.15*** (2.73)	0.01	0.02	2,424
Stake	0.01 (0.49)	0.06 (0.50)	0.07	0.08	2,424
Apply sand	0.01* (1.70)	0.18 (1.62)	0.06	0.07	2,424
Use trap crops	0.01 (1.30)	0.07 (1.24)	0.01	0.02	2,424
Burn crop residue for pest/disease control	0.00 (0.06)	0.01 (0.06)	0.06	0.06	2,424
Use traps	-0.00 (-0.07)	-0.00 (-0.07)	0.02	0.02	2,424
Cover with leaves	0.00 (0.54)	0.03 (0.54)	0.01	0.01	2,424

Note. Robust *t* statistics clustered at the clinic level in parentheses. All estimations control for lottery fixed effects. Impact estimates are 36-month impact estimates. Means are 36-month means. ITT = Intention to Treat. LATE=Local Average Treatment Effect, *p < .10; **p < .05; ***p < .01

Table A1.3 Impacts on Cultural Practices: Perennial Crops (Single Difference)

Dependent variable	ITT impact (1)	LATE impact (2)	Control mean (3)	Treatment mean (4)	N (5)
Organic fertilizer used	-0.05	-0.49	0.60	0.56	1,377
	(-1.58)	(-1.51)			
Inorganic fertilizer used	-0.01	-0.11	0.27	0.25	1,377
	(-0.45)	(-0.46)			
Pesticide used	-0.03	-0.28	0.25	0.22	1,377
	(-1.11)	(-1.09)			
Use resistant variety	0.04	0.44	0.36	0.40	1,377
	(1.49)	(1.44)			
Improved planting material	-0.02	-0.19	0.28	0.26	1,377
	(-0.69)	(-0.68)			
Certified planting material	-0.04	-0.37	0.25	0.22	1,377
	(-1.41)	(-1.34)			
Intercrop	0.00	0.01	0.21	0.19	1,377
	(0.07)	(0.07)			
Weed around crop	0.02	0.17	0.79	0.80	1,377
	(0.77)	(0.77)			
Remove infested/damaged material	0.06*	0.65*	0.36	0.43	1,377
	(1.96)	(1.95)			
Mulch	0.02	0.17	0.08	0.09	1,377
	(1.05)	(1.05)			
Apply ash	0.01	0.15	0.07	0.09	1,377
	(1.08)	(1.09)			
Spray with chilies?	0.00	0.03	0.00	0.01	1,377
	(0.85)	(0.87)			
Prune	-0.01	-0.06	0.47	0.46	1,377
	(-0.22)	(-0.22)			
Change cycle	-0.00	-0.00	0.03	0.03	1,377
	(-0.04)	(-0.04)			
Use trap crops	-0.00	-0.02	0.00	0.00	1,377
	(-0.67)	(-0.66)			
Burn crop residue for pest/disease control	-0.00	-0.03	0.02	0.02	1,377
	(-0.56)	(-0.57)			
Use traps	-0.00	-0.04	0.01	0.01	1,377
	(-0.77)	(-0.77)			
Grease	0.01**	0.06**	0.00	0.01	1,377
	(2.32)	(2.35)			

Note. Robust *t* statistics clustered at the clinic level in parentheses. All estimations control for lottery fixed effects. Impact estimates are 36-month impact estimates. Means are 36-month means. ITT = Intention to Treat. LATE=Local Average Treatment Effect, *p < .10; **p < .05; ***p < .01

Table A1.4 Impacts on Crop Diversity: All Crops (Single Difference)

Dependent variable	ITT impact (1)	LATE impact (2)	Control mean (3)	Treatment mean (4)	N (5)
No. of all crops produced	-0.23	-2.75	9.10	8.95	2,550
	(-0.67)	(-0.66)			
No. crops produced large area	-0.10	-1.15	3.94	3.86	2,550
	(-0.91)	(-0.89)			
Production area in ha	-0.05	-0.58	3.03	3.03	2,550
	(-0.36)	(-0.36)			
Log(value of production per ha)	-0.04	-0.52	9.37	9.33	2,226
	(-0.42)	(-0.42)			
Log(value of costs per ha)	0.01	0.16	8.84	8.87	2,542
	(0.17)	(0.18)			
Log(gross margins per ha)	0.18	2.21	3.75	3.92	2,542
	(0.51)	(0.52)			

Note. Robust *t* statistics clustered at the clinic level in parentheses. All estimations control for lottery fixed effects. Impact estimates are 36-month impact estimates. Means are 36-month means. ITT = Intention to Treat. LATE=Local Average Treatment Effect, *p < .10; **p < .05; ***p < .01

Table A1.5 Impacts on Crop Diversity: Annual Crops (Single Difference)

Dependent variable	ITT impact (1)	LATE impact (2)	Control mean (3)	Treatment mean (4)	N (5)
No. of all crops produced	-0.13	-1.56	4.00	3.88	2,424
	(-1.19)	(-1.15)			
No. crops produced large area	-0.07	-0.78	2.09	2.02	2,424
	(-1.50)	(-1.37)			
Production area in ha	-0.21*	-2.49*	2.70	2.53	2,424
	(-1.84)	(-1.76)			

Note. Robust *t* statistics clustered at the clinic level in parentheses. All estimations control for lottery fixed effects. Impact estimates are 36-month impact estimates. Means are 36-month means. ITT = Intention to Treat. LATE=Local Average Treatment Effect, *p < .10; **p < .05; ***p < .01

Table A1.6 Impacts on Crop Diversity: Perennial Crops (Single Difference)

Dependent variable	ITT impact (1)	LATE impact (2)	Control mean (3)	Treatment mean (4)	N (5)
No. of all crops produced	-0.14	-1.44	5.29	5.07	1,377
	(-1.00)	(-0.98)			
No. crops produced large area	0.01	0.08	1.51	1.51	1,377
	(0.32)	(0.32)			
Production area in ha	0.25**	2.56**	0.85	1.11	1,377
	(2.61)	(2.16)			

Note. Robust *t* statistics clustered at the clinic level in parentheses. All estimations control for lottery fixed effects. Impact estimates are 36-month impact estimates. Means are 36-month means. ITT = Intention to Treat. LATE=Local Average Treatment Effect, *p < .10; **p < .05; ***p < .01

Table A1.7 Impacts on Production: Annual Crops (Single Difference)

Dependent variable	ITT impact (1)	LATE impact (2)	Control mean (3)	Treatment mean (4)	N (5)
Value of production per acre	0.01	0.13	9.27	9.30	1,951
	(0.11)	(0.11)			
Cost of seed planted per acre	0.10	1.27	7.71	7.82	1,951
	(1.01)	(1.04)			
Cost of inorganic fertilizer per acre	0.01	0.09	6.51	6.63	1,951
	(0.05)	(0.05)			
Cost of pesticide per acre	-0.08	-1.06	3.20	3.24	1,951
	(-0.51)	(-0.50)			
Cost of labor per acre	0.20	2.58	4.53	4.79	1,951
	(1.10)	(1.15)			
Total costs per acre	0.10	1.33	9.08	9.20	1,951
	(1.51)	(1.54)			
Gross margins per acre	-0.40	-5.07	3.09	2.57	1,951
	(-0.93)	(-0.92)			

Note. All outcome values are in natural logs. Robust *t* statistics clustered at the clinic level in parentheses. All estimations control for lottery fixed effects. Impact estimates are 36-month impact estimates. Means are 36-month means. ITT = Intention to Treat. LATE=Local Average Treatment Effect, *p < .10; **p < .05; ***p < .01.

Table A1.8 Impacts on Production: Perennial Crops (Single Difference)

Dependent variable	ITT impact (1)	LATE impact (2)	Control mean (3)	Treatment mean (4)	N (5)
Value of production per acre	-0.28	-2.49	9.87	9.52	1,011
	(-1.57)	(-1.44)			
Cost of inorganic fertilizer per acre	-0.17	-1.51	2.84	2.51	1,011
	(-0.73)	(-0.65))			
Cost of pesticide per acre	-0.36	-3.21	2.60	2.17	1,011
	(-1.27)	(-1.18)			
Cost of labor per acre	0.37	3.31	2.25	2.41	1,011
	(1.45)	(1.42)			
Total costs per acre	0.42	3.79	4.24	4.46	1,011
	(1.38)	(1.36)			
Gross margins per acre	0.05	0.49	7.76	7.79	1,011
	(0.15)	(0.15)			

Note. All outcome values are in natural logs. Robust *t* statistics clustered at the clinic level in parentheses. All estimations control for lottery fixed effects. Impact estimates are 36-month impact estimates. Means are 36-month means. ITT = Intention to Treat. LATE=Local Average Treatment Effect, *p < .10; **p < .05; ***p < .01.

Table A1.9 Impacts on Cultural Practices and Input Use: Maize Producers (Single Difference)

Dependent variable	ITT impact (1)	LATE impact (2)	Control mean (3)	Treatment mean (4)	N (5)
Organic fertilizer used	-0.02	-0.23	0.45	0.42	2,161
	(-0.99)	(-0.98)			
No. organic fertilizer applications	0.04	0.33	1.24	1.29	921
	(1.25)	(1.29)			
Inorganic fertilizer used	-0.01	-0.06	0.82	0.82	2,161
	(-0.32)	(-0.32)			
No. inorganic fertilizer applications	-0.02	-0.17	1.18	1.15	1,755
	(-0.77)	(-0.78)			
Family labor days (fertilizer)	-0.06	-0.55	2.99	2.98	1,670
	(-0.41)	(-0.41)			
Paid labor days (fertilizer)	-0.15	-1.64	3.09	2.92	687
	(-0.60)	(-0.57)			
All labor days (fertilizer)	0.04	0.35	3.75	3.81	1,887
	(0.19)	(0.20)			
Pesticide used	-0.02	-0.17	0.23	0.21	2,161
	(-0.81)	(-0.81)			
No. pesticide applications	0.00	0.02	1.36	1.37	477
	(0.05)	(0.05)			
Family labor days (pesticide)	7.32	74.92	1.85	9.09	344
	(1.42)	(1.23)			
Paid labor days (pesticide)	4.29	27.16	7.33	10.07	222
	(0.44)	(0.38)			
All labor days (pesticide)	11.57	97.08	4.35	17.10	523
	(1.38)	(1.22)			

Crop rotation	0.06***	0.70**	0.20	0.28	2,160
	(3.03)	(2.63)			
Resistant Variety	0.02	0.22	0.44	0.47	2,160
	(1.12)	(1.09)			
Used improved planting material	0.01	0.07	0.77	0.79	2,160
	(0.27)	(0.27)			
Used certified planting material	-0.01	-0.08	0.85	0.85	2,160
	(-0.39)	(-0.40)			
Remove plant residue from prior harvest	0.02	0.19	0.91	0.93	2,160
	(1.32)	(1.24)			
Plant Early	0.03	0.33	0.71	0.74	2,160
	(1.43)	(1.43)			
Intercrop	-0.08***	-0.92***	0.71	0.62	2,160
	(-3.62)	(-2.85)			
Times check crop	0.54	5.89	12.37	13.27	2,160
	(0.73)	(0.73)			
Weed in a timely manner	0.01	0.07	0.95	0.96	2,160
	(0.77)	(0.78)			
Remove volunteer crops	0.07***	0.72**	0.58	0.65	2,160
	(2.99)	(2.61)			
Remove infested or damaged material	0.03	0.30	0.56	0.59	2,160
	(1.18)	(1.21)			
Mulch	-0.01	-0.11	0.03	0.02	2,160
	(-1.43)	(-1.44)			
Apply ash	0.02*	0.23*	0.04	0.06	2,160
	(1.85)	(1.80)			
Spray with chillies	0.01**	0.10**	0.01	0.02	2,160
	(2.53)	(2.46)			
Stake	0.01	0.12	0.04	0.05	2,160
	(1.10)	(1.09)			
Apply sand	0.02**	0.23**	0.04	0.06	2,160
	(2.38)	(2.05)			
Use trap crops	0.00	0.05	0.01	0.02	2,160
	(1.13)	(1.13)			
Burn crop residue pest/diseases control	0.01	0.12	0.03	0.05	2,160
	(1.30)	(1.33)			
Use traps	-0.00	-0.04	0.02	0.01	2,160
	(-0.71)	(-0.70)			
Cover with leaves	-0.00	-0.02	0.01	0.01	2,160
	(-0.61)	(-0.61)			

Note. Robust *t* statistics clustered at the clinic level in parentheses. All estimations control for lottery fixed effects. Impact estimates are 36-month impact estimates. Means are 36-month means. ITT = Intention to Treat. LATE=Local Average Treatment Effect, *p < .10; **p < .05; ***p < .01.

Table A1.10 Impacts on Cultural Practices and Input Use: Maize Producers (ANCOVA)

Dependent variable	ITT impact (1)	LATE impact (2)	Control mean (3)	Treatment mean (4)	N (5)
Organic fertilizer used	-0.03	-0.34	0.42	0.40	1,443
	(-1.30)	(-1.25)			
No. organic fertilizer applications	0.03	0.19	1.28	1.30	592
	(0.57)	(0.59)			
Inorganic fertilizer used	-0.00	-0.02	0.84	0.85	1,443
	(-0.10)	(-0.10)			
No. inorganic fertilizer applications	-0.05	-0.48	1.22	1.16	1,226
	(-1.49)	(-1.47)			
Family labor days (fertilizer)	-0.15	-1.47	3.09	2.99	1,147
	(-0.94)	(-0.91)			
Paid labor days (fertilizer)	-0.13	-1.25	3.01	2.94	503
	(-0.39)	(-0.38)			
All labor days (fertilizer)	-0.04	-0.45	3.89	3.88	1,311
	(-0.20)	(-0.20)			
Pesticide used	-0.03	-0.28	0.25	0.22	1,443
	(-1.11)	(-1.07)			
No. pesticide applications	0.04	0.37	1.35	1.41	343
	(0.85)	(0.78)			
Family labor days (pesticide)	6.27	57.68	1.94	7.42	241
	(1.05)	(0.99)			
Paid labor days (pesticide)	-0.28	-1.80	2.13	1.93	175
	(-1.05)	(-1.01)			
All labor days (pesticide)	12.37	108.35	2.24	14.95	379
	(1.16)	(1.05)			
Crop rotation	0.05**	0.60*	0.21	0.28	1,428
	(2.10)	(1.92)			
Resistant Variety	0.04*	0.48*	0.46	0.51	1,428
	(1.93)	(1.74)			
Used improved planting material	-0.01	-0.14	0.83	0.83	1,428
	(-0.43)	(-0.43)			
Used certified planting material	-0.00	-0.00	0.86	0.87	1,428
	(-0.00)	(-0.00)			
Remove plant residue from prior harvest	0.01	0.06	0.92	0.94	1,428
	(0.31)	(0.32)			
Plant Early	0.00	0.02	0.78	0.78	1,428
	(0.06)	(0.06)			
Intercrop	-0.09***	-0.96**	0.72	0.62	1,428
	(-2.78)	(-2.33)			
Times check crop	0.09	1.04	14.02	14.43	1,428
	(0.10)	(0.10)			
Weed in a timely manner	0.00	0.05	0.97	0.97	1,428
	(0.52)	(0.53)			
Remove volunteer crops	0.07**	0.77**	0.56	0.64	1,428
	(2.46)	(2.09)			
Remove infested or damaged material	0.02	0.23	0.59	0.61	1,428

	(0.72)	(0.74)			
Mulch	-0.01	-0.15	0.03	0.02	1,428
	(-1.43)	(-1.45)			
Apply ash	0.02*	0.19*	0.04	0.06	1,428
	(1.78)	(1.71)			
Spray with chillies	0.02***	0.18**	0.01	0.02	1,428
	(2.65)	(2.35)			
Stake	0.02	0.19	0.05	0.07	1,428
	(1.12)	(1.09)			
Apply sand	0.02*	0.26	0.06	0.08	1,428
	(1.75)	(1.51)			
Use trap crops	0.00	0.03	0.01	0.02	1,428
	(0.70)	(0.71)			
Burn crop residue pest/diseases control	0.01	0.07	0.04	0.05	1,428
	(0.59)	(0.60)			
Use traps	-0.01*	-0.12	0.02	0.01	1,428
	(-1.69)	(-1.61)			
Cover with leaves	-0.00	-0.03	0.01	0.01	1,428
	(-0.68)	(-0.69)			

Note. Robust *t* statistics clustered at the clinic level in parentheses. All estimations control for lottery fixed effects. Impact estimates are 36-month impact estimates. Means are 36-month means. ITT = Intention to Treat. LATE=Local Average Treatment Effect, *p < .10; **p < .05; ***p < .01

Table A1.11 Pesticide knowledge and practice: Maize Producers (Single Difference)

Dependent variable	ITT impact (1)	Control mean (2)	Treatment mean (3)	N (4)
Pre-harvest interval is important	-0.01	0.46	0.46	2,233
	(-0.23)			
Check for plant health problems on a regular basis	0.05**	0.72	0.77	2,233
	(2.38)			
Prefers chemical pest control	-0.06**	0.62	0.57	2,233
	(-2.45)			
No. protective items for pesticide app	-0.26**	2.34	2.13	2,233
	(-2.52)			
Spray pesticide in the morning	-0.04**	0.82	0.79	1,629
	(-2.23)			
Spray pesticide in the evening	-0.01	0.29	0.28	1,629
	(-0.56)			
Avoid chemical drift when spraying	0.09***	0.27	0.36	1,629
	(3.79)			
Washing self after spraying	0.02	0.61	0.63	1,629
	(0.82)			
Washing equipment after spraying	0.01	0.57	0.57	1,629
	(0.24)			
Using containers only for pesticide	0.00	0.43	0.42	1,629
	(0.05)			
Chemical disposal: use it all	-0.00	0.26	0.26	1,629
	(-0.21)			
Chemical storage: shed or barn	0.00	0.46	0.47	1,640
	(0.21)			

Note. Robust t statistics clustered at the clinic level in parentheses. All estimations control for lottery fixed effects. Impact estimates are 36-month impact estimates. Means are 36-month means. ITT = Intention to Treat. LATE=Local Average Treatment Effect, *p < .10; **p < .05; ***p < .01.

Table A1.12 Impacts on Production: Maize Producers (Single Difference)

Dependent variable	ITT impact (1)	LATE impact (2)	Control mean (3)	Treatment mean (4)	N (5)
Value of production per acre	0.13**	1.40**	9.42	9.55	1,460
	(2.03)	(2.04)			
Cost of seed planted per acre	0.04	0.39	7.69	7.73	1,460
	(1.23)	(1.26)			
Cost of inorganic fertilizer per acre	-0.11	-1.16	6.88	6.82	1,460
	(-0.62)	(-0.61)			
Cost of pesticide per acre	-0.18	-1.84	1.86	1.68	1,460
	(-1.06)	(-1.04)			
Cost of labor per acre	-0.05	-0.54	4.31	4.24	1,460
	(-0.23)	(-0.23)			
Total costs per acre)	0.03	0.31	8.94	8.97	1,460
	(0.57)	(0.59)			
Gross margins per acre	0.13	1.33	5.08	5.03	1,458
	(0.35)	(0.36)			

Note. All outcome values are in natural logs. Robust *t* statistics clustered at the clinic level in parentheses. All estimations control for lottery fixed effects. Impact estimates are 36-month impact estimates. Means are 36-month means. ITT = Intention to Treat. LATE=Local Average Treatment Effect, *p < .10; **p < .05; ***p < .01.

Table A1.13 Impacts on Cultural Practices and Input Use: Bean Producers (Single Difference)

Dependent variable	ITT impact (1)	LATE impact (2)	Control mean (3)	Treatment mean (4)	N (5)
Organic fertilizer used	-0.02	-0.25	0.39	0.37	1,504
	(-0.94)	(-0.93)			
No. organic fertilizer applications	0.07	0.44*	1.20	1.26	554
	(1.61)	(1.68)			
Inorganic fertilizer used	-0.02	-0.20	0.56	0.54	1,504
	(-0.78)	(-0.80)			
No. inorganic fertilizer applications	0.02	0.18	1.14	1.12	808
	(0.79)	(0.79)			
Family labor days (fertilizer)	-0.11	-1.00	2.97	2.84	890
	(-0.58)	(-0.60)			
Paid labor days (fertilizer)	-0.08	-1.03	3.17	2.71	353
	(-0.24)	(-0.24)			
All labor days (fertilizer)	-0.03	-0.24	3.85	3.71	971
	(-0.11)	(-0.11)			
Pesticide used	-0.03*	-0.35*	0.17	0.14	1,504
	(-1.75)	(-1.77)			
No. pesticide applications	-0.15**	-1.30	1.46	1.44	235
	(-2.33)	(-1.41)			
Family labor days (pesticide)	21.56	130.52	9.69	16.71	178
	(1.56)	(1.46)			
Paid labor days (pesticide)	-3.52	-32.00	28.02	34.72	125
	(-0.57)	(-0.56)			
All labor days (pesticide)	15.25	97.21	18.81	20.16	272
	(1.56)	(1.55)			
Crop rotation	0.03	0.28	0.22	0.25	1,504
	(1.08)	(1.07)			
Resistant Variety	-0.00	-0.03	0.21	0.22	1,504
	(-0.13)	(-0.13)			
Used improved planting material	-0.00	-0.01	0.35	0.35	1,504
	(-0.03)	(-0.03)			
Used certified planting material	-0.01	-0.09	0.26	0.26	1,504
	(-0.36)	(-0.36)			
Remove plant residue from prior harvest	0.01	0.06	0.81	0.81	1,504
	(0.30)	(0.30)			
Plant Early	0.02	0.21	0.69	0.72	1,504
	(0.84)	(0.84)			
Intercrop	-0.02	-0.20	0.87	0.85	1,504
	(-1.07)	(-1.07)			
Times check crop	0.99	10.33	11.24	12.53	1,504
	(1.33)	(1.34)			
Weed in a timely manner	0.01	0.06	0.93	0.94	1,504
	(0.43)	(0.44)			
Remove volunteer crops	0.06**	0.66**	0.49	0.56	1,504
	(2.23)	(2.12)			
Remove infested or damaged material	0.05*	0.56*	0.44	0.49	1,504

	(1.86)	(1.90)			
Mulch	-0.00	-0.05	0.03	0.03	1,504
	(-0.61)	(-0.62)			
Apply ash	0.02*	0.16*	0.02	0.04	1,504
	(1.72)	(1.71)			
Spray with chillies	0.01	0.07	0.01	0.01	1,504
	(1.60)	(1.52)			
Stake	0.01**	0.10**	0.00	0.01	1,504
	(2.49)	(2.29)			
Apply sand	0.00	0.04	0.04	0.04	1,504
	(0.32)	(0.32)			
Use trap crops	0.01*	0.08*	0.00	0.01	1,504
	(1.87)	(1.75)			
Burn crop residue pest/diseases control	-0.01	-0.08	0.05	0.04	1,504
	(-0.93)	(-0.91)			
Use traps	0.00	0.02	0.01	0.01	1,504
	(0.68)	(0.70)			
Cover with leaves	0.00	0.03	0.00	0.01	1,504
	(1.09)	(1.08)			

Note. Robust *t* statistics clustered at the clinic level in parentheses. All estimations control for lottery fixed effects. Impact estimates are 36-month impact estimates. Means are 36-month means. ITT = Intention to Treat. LATE=Local Average Treatment Effect, *p < .10; **p < .05; ***p < .01

Table A1.14 Impacts on Cultural Practices and Input Use: Bean Producers (ANCOVA)

Dependent variable	ITT impact (1)	LATE impact (2)	Control mean (3)	Treatment mean (4)	N (5)
Organic fertilizer used	-0.04	-0.36	0.38	0.36	1,134
	(-1.31)	(-1.30)			
No. organic fertilizer applications	0.08	0.46*	1.19	1.27	414
	(1.65)	(1.78)			
Inorganic fertilizer used	-0.02	-0.20	0.56	0.53	1,134
	(-0.87)	(-0.88)			
No. inorganic fertilizer applications	0.00	0.03	1.13	1.11	618
	(0.12)	(0.12)			
Family labor days (fertilizer)	-0.18	-1.60	2.97	2.87	683
	(-0.84)	(-0.86)			
Paid labor days (fertilizer)	-0.03	-0.28	2.81	2.65	285
	(-0.07)	(-0.08)			
All labor days (fertilizer)	-0.10	-0.95	3.78	3.78	749
	(-0.46)	(-0.46)			
Pesticide used	-0.04*	-0.35	0.17	0.15	1,134
	(-1.77)	(-1.64)			
No. pesticide applications	-0.11	-0.62	1.39	1.44	181
	(-1.43)	(-1.35)			
Family labor days (pesticide)	28.25	138.59	11.53	20.64	142
	(1.56)	(1.44)			
Paid labor days (pesticide)	-23.79	-196.95	39.16	33.68	101
	(-1.51)	(-1.25)			

All labor days (pesticide)	12.16	69.26	24.21	21.43	215
	(0.87)	(0.92)			
Crop rotation	0.01	0.09	0.23	0.25	1,121
	(0.39)	(0.40)			
Resistant Variety	0.02	0.16	0.20	0.24	1,121
	(0.66)	(0.66)			
Used improved planting material	-0.01	-0.11	0.39	0.39	1,121
	(-0.38)	(-0.39)			
Used certified planting material	-0.02	-0.22	0.28	0.26	1,121
	(-0.90)	(-0.90)			
Remove plant residue from prior harvest	0.00	0.04	0.79	0.80	1,121
	(0.18)	(0.18)			
Plant Early	-0.02	-0.19	0.72	0.73	1,121
	(-0.68)	(-0.68)			
Intercrop	-0.01	-0.08	0.86	0.85	1,121
	(-0.41)	(-0.42)			
Times check crop	0.16	1.60	12.06	13.12	1,121
	(0.15)	(0.16)			
Weed in a timely manner	-0.00	-0.01	0.93	0.94	1,121
	(-0.10)	(-0.10)			
Remove volunteer crops	0.09***	0.93**	0.47	0.56	1,121
	(2.65)	(2.40)			
Remove infested or damaged material	0.05	0.53	0.44	0.48	1,121
	(1.42)	(1.47)			
Mulch	-0.01	-0.07	0.02	0.02	1,121
	(-0.86)	(-0.87)			
Apply ash	0.01	0.12	0.03	0.04	1,121
	(1.01)	(1.05)			
Spray with chillies	0.01	0.09	0.01	0.01	1,121
	(1.65)	(1.60)			
Stake	0.01*	0.09*	0.01	0.01	1,121
	(1.90)	(1.79)			
Apply sand	0.01	0.06	0.05	0.05	1,121
	(0.38)	(0.38)			
Use trap crops	0.01	0.06	0.00	0.01	1,121
	(1.22)	(1.27)			
Burn crop residue pest/diseases control	-0.02*	-0.19*	0.05	0.04	1,121
	(-1.99)	(-1.82)			
Use traps	-0.00	-0.01	0.01	0.00	1,121
	(-0.16)	(-0.17)			
Cover with leaves	0.00	0.03	0.00	0.00	1,121
	(1.29)	(1.32)			

Note. Robust t statistics clustered at the clinic level in parentheses. All estimations control for lottery fixed effects. Impact estimates are 36-month impact estimates. Means are 36-month means. ITT = Intention to Treat. LATE=Local Average Treatment Effect, *p < .10; **p < .05; ***p < .01

Table A1.15. Pesticide knowledge and practice: Bean Producers Only (Single Difference)

Dependent variable	ITT impact (1)	Control mean (2)	Treatment mean (3)	N (4)
Pre-harvest interval is important	0.00	0.46	0.46	1,602
	(0.03)			
Check for plant health problems on a regular basis	0.06**	0.71	0.77	1,602
	(2.19)			
Prefers chemical pest control	-0.07**	0.63	0.56	1,602
	(-2.43)			
No. protective items for pesticide app	-0.27**	2.43	2.18	1,602
	(-2.23)			
Spray pesticide in the morning	-0.05**	0.82	0.79	1,175
	(-2.07)			
Spray pesticide in the evening	-0.03	0.29	0.27	1,175
	(-1.25)			
Avoid chemical drift when spraying	0.11***	0.25	0.35	1,175
	(3.94)			
Washing self after spraying	0.05*	0.60	0.63	1,175
	(1.91)			
Washing equipment after spraying	0.06**	0.55	0.59	1,175
	(2.27)			
Using containers only for pesticide	0.01	0.42	0.42	1,175
	(0.19)			
Chemical disposal: use it all	-0.01	0.28	0.27	1,175
	(-0.40)			
Chemical storage: shed or barn	0.00	0.44	0.46	1,182
	(0.19)			

Note. Robust t statistics clustered at the clinic level in parentheses. All estimations control for lottery fixed effects. Impact estimates are 36-month impact estimates. Means are 36-month means. ITT = Intention to Treat. LATE=Local Average Treatment Effect, *p < .10; **p < .05; ***p < .01

Table A1.16 Impacts on Production: Bean Producers (Single Difference)

Dependent variable	RF impact (1)	IV impact (2)	Control mean (3)	Treatment mean (4)	N (5)
Value of production per acre	-0.09	-0.89	8.61	8.51	1,135
	(-1.42)	(-1.29)			
Cost of seed planted per acre	-0.12	-1.16	7.11	7.02	1,135
	(-1.15)	(-1.43)			
Cost of inorganic fertilizer per acre	-0.09	-0.89	4.30	4.19	1,135
	(-0.46)	(-0.47)			
Cost of pesticide per acre	-0.11	-1.07	1.10	1.05	1,135
	(-0.81)	(-0.80)			
Cost of labor per acre	-0.10	-0.91	3.82	3.83	1,135
	(-0.37)	(-0.37)			
Total costs per acre)	-0.11	-1.05	8.35	8.25	1,135
	(-1.63)	(-1.58)			
Gross margins per acre	-0.61	-5.52	1.82	1.12	1,134
	(-1.11)	(-1.05)			

Note. All outcome values are in natural logs. Robust *t* statistics clustered at the clinic level in parentheses. All estimations control for lottery fixed effects. Impact estimates are 36-month impact estimates. Means are 36-month means. ITT = Intention to Treat. LATE=Local Average Treatment Effect, *p < .10; **p < .05; ***p < .01.

Table A1.17 Impacts on Cultural Practices and Input Use: Potato Producers (Single Difference)

Dependent variable	ITT impact (1)	LATE impact (2)	Control mean (3)	Treatment mean (4)	N (5)
Organic fertilizer used	0.07	1.10	0.21	0.25	441
	(1.29)	(0.92)			
No. organic fertilizer applications	-0.02	-0.65	1.19	1.21	104
	(-0.68)	(-0.55)			
Inorganic fertilizer used	0.00	0.04	0.71	0.79	441
	(0.07)	(0.07)			
No. inorganic fertilizer applications	-0.07	-0.77	1.49	1.47	335
	(-0.38)	(-0.39)			
Family labor days (fertilizer)	0.19	2.36	4.37	4.89	323
	(0.28)	(0.29)			
Paid labor days (fertilizer)	-0.97	-30.75	6.38	5.89	158
	(-1.31)	(-0.55)			
All labor days (fertilizer)	0.09	1.62	6.32	6.76	366
	(0.15)	(0.16)			
Pesticide used	0.02	0.25	0.54	0.65	441
	(0.53)	(0.52)			
No. pesticide applications	-0.02	-0.20	2.81	2.78	268
	(-0.06)	(-0.06)			
Family labor days (pesticide)	-0.43	-3.93	4.09	3.52	175
	(-0.79)	(-0.85)			
Paid labor days (pesticide)	6.14	88.14	2.96	12.39	119
	(0.95)	(0.87)			
All labor days (pesticide)	1.79	19.65	4.03	8.16	268
	(0.93)	(1.00)			
Crop rotation	0.17***	2.62*	0.26	0.47	435
	(2.76)	(1.68)			
Resistant Variety	-0.04	-0.54	0.22	0.23	435
	(-0.74)	(-0.83)			
Used improved planting material	0.05	0.74	0.40	0.47	435
	(1.06)	(1.06)			
Used certified planting material	0.00	0.08	0.28	0.29	435
	(0.10)	(0.10)			
Remove plant residue from prior harvest	-0.01	-0.22	0.83	0.85	435
	(-0.47)	(-0.43)			
Plant Early	0.05	0.75	0.64	0.68	435
	(1.24)	(0.98)			
Intercrop	-0.04	-0.55	0.21	0.13	435
	(-1.27)	(-1.56)			
Times check crop	-1.45	-22.58	12.06	11.95	435
	(-0.85)	(-0.87)			
Weed in a timely manner	-0.03	-0.47	0.96	0.93	435
	(-1.39)	(-1.47)			
Remove volunteer crops	0.03	0.42	0.59	0.68	435
	(0.36)	(0.34)			
Remove infested or damaged material	0.04	0.58	0.55	0.61	435

	(1.23)	(0.95)			
Mulch	-0.00	-0.05	0.10	0.06	435
	(-0.25)	(-0.24)			
Apply ash	0.01	0.23	0.01	0.02	435
	(1.29)	(1.07)			
Spray with chillies	0.00*	-0.00	0.00	0.00	435
	(1.83)	(-1.39)			
Stake	0.00	0.07	0.00	0.00	435
	(0.94)	(0.69)			
Apply sand	-0.01	-0.17	0.14	0.16	435
	(-0.33)	(-0.39)			
Use trap crops	0.02**	0.36	0.00	0.02	435
	(2.19)	(1.64)			
Burn crop residue pest/diseases control	0.02**	0.30	0.00	0.02	435
	(2.05)	(1.30)			
Use traps	0.01	0.08	0.00	0.01	435
	(1.40)	(1.36)			
Cover with leaves	0.00	0.06	0.00	0.01	435
	(1.35)	(1.22)			

Note. Robust t statistics clustered at the clinic level in parentheses. All estimations control for lottery fixed effects. Impact estimates are 36-month impact estimates. Means are 36-month means. ITT = Intention to Treat. LATE=Local Average Treatment Effect, *p < .10; **p < .05; ***p < .01

Table A1.18 Impacts on Cultural Practices and Input Use: Potato Producers (ANCOVA)

Dependent variable	ITT impact (1)	LATE impact (2)	Control mean (3)	Treatment mean (4)	N (5)
Organic fertilizer used	0.08	1.17	0.21	0.25	440
	(1.43)	(0.99)			
No. organic fertilizer applications	-0.01	-0.17	1.19	1.21	104
	(-0.14)	(-0.16)			
Inorganic fertilizer used	0.00	0.05	0.71	0.80	440
	(0.09)	(0.10)			
No. inorganic fertilizer applications	-0.07	-0.77	1.49	1.47	335
	(-0.38)	(-0.40)			
Family labor days (fertilizer)	0.23	2.78	4.37	4.89	323
	(0.32)	(0.32)			
Paid labor days (fertilizer)	-0.91	-28.30	6.38	5.89	158
	(-1.31)	(-0.54)			
All labor days (fertilizer)	0.12	2.22	6.32	6.76	366
	(0.20)	(0.21)			
Pesticide used	0.01	0.23	0.54	0.66	440
	(0.47)	(0.47)			
No. pesticide applications	-0.02	-0.24	2.81	2.78	268
	(-0.07)	(-0.07)			
Family labor days (pesticide)	-0.43	-3.91	4.09	3.52	175
	(-0.77)	(-0.84)			
Paid labor days (pesticide)	6.38	96.00	2.96	12.39	119
	(0.94)	(0.86)			

All labor days (pesticide)	1.52	17.09	4.03	8.16	268
	(0.74)	(0.77)			
Crop rotation	0.16***	2.62	0.26	0.47	433
	(2.87)	(1.56)			
Resistant Variety	-0.03	-0.46	0.22	0.23	433
	(-0.61)	(-0.68)			
Used improved planting material	0.05	0.79	0.40	0.47	433
	(1.07)	(1.05)			
Used certified planting material	0.01	0.11	0.28	0.29	433
	(0.14)	(0.14)			
Remove plant residue from prior harvest	-0.01	-0.16	0.83	0.85	433
	(-0.32)	(-0.31)			
Plant Early	0.04	0.68	0.64	0.67	433
	(1.16)	(0.92)			
Intercrop	-0.02	-0.26	0.20	0.13	433
	(-0.58)	(-0.71)			
Times check crop	-1.64	-26.32	12.10	11.98	433
	(-0.91)	(-0.92)			
Weed in a timely manner	-0.03	-0.47	0.96	0.93	433
	(-1.40)	(-1.47)			
Remove volunteer crops	0.02	0.38	0.59	0.68	433
	(0.31)	(0.30)			
Remove infested or damaged material	0.03	0.45	0.56	0.61	433
	(0.89)	(0.76)			
Mulch	-0.01	-0.12	0.10	0.06	433
	(-0.55)	(-0.51)			
Apply ash	0.02	0.24	0.01	0.02	433
	(1.19)	(1.00)			
Spray with chillies	0.00***	0.00	0.00	0.00	433
	(3.75)	(1.36)			
Stake	0.00	0.07	0.00	0.00	433
	(0.94)	(0.70)			
Apply sand	-0.01	-0.18	0.15	0.16	433
	(-0.33)	(-0.39)			
Use trap crops	0.02**	0.38	0.00	0.02	433
	(2.19)	(1.62)			
Burn crop residue pest/diseases control	0.02**	0.32	0.00	0.02	433
	(2.07)	(1.27)			
Use traps	0.01	0.09	0.00	0.01	433
	(1.39)	(1.36)			
Cover with leaves	0.00	0.07	0.00	0.01	433
	(1.37)	(1.22)			

Note. Robust t statistics clustered at the clinic level in parentheses. All estimations control for lottery fixed effects. Impact estimates are 36-month impact estimates. Means are 36-month means. ITT = Intention to Treat. LATE=Local Average Treatment Effect, *p < .10; **p < .05; ***p < .01

Table A1.9 Impacts on Production: Potato Producers (Single Difference)

Dependent variable	ITT impact (1)	LATE impact (2)	Control mean (3)	Treatment mean (4)	N (5)
Value of production per acre	-0.07	-1.17	10.37	10.32	359
	(-0.62)	(-0.58)			
Cost of seed planted per acre	-0.65	-11.01	9.65	9.08	359
	(-1.62)	(-1.41)			
Cost of inorganic fertilizer per acre	0.01	0.20	6.66	7.34	359
	(0.04)	(0.05)			
Cost of pesticide per acre	-0.59*	-10.03	5.01	5.21	359
	(-1.81)	(-1.58)			
Cost of labor per acre	0.56	9.52	4.93	5.90	359
	(1.13)	(1.50)			
Total costs per acre)	-0.15	-2.49	11.37	11.37	359
	(-1.18)	(-1.03)			
Gross margins per acre	1.46	28.23	-7.75	-6.43	350
	(1.63)	(1.51)			

Note. All outcome values are in natural logs. Robust *t* statistics clustered at the clinic level in parentheses. All estimations control for lottery fixed effects. Impact estimates are 36-month impact estimates. Means are 36-month means. ITT = Intention to Treat. LATE=Local Average Treatment Effect, *p < .10; **p < .05; ***p < .01.

Table A1.20 Impacts on Cultural Practices and Input Use: Coffee Producers (Single Difference)

Dependent variable	ITT impact (1)	LATE impact (2)	Control mean (3)	Treatment mean (4)	N (5)
Organic fertilizer used	-0.04	-0.37	0.65	0.61	403
	(-0.96)	(-0.93)			
Inorganic fertilizer used	-0.06	-0.53	0.63	0.57	403
	(-1.15)	(-1.14)			
Pesticide used	-0.08**	-0.75	0.59	0.53	403
	(-2.04)	(-1.68)			
Resistant variety	-0.08*	-0.78	0.52	0.48	401
	(-1.69)	(-1.54)			
Used improved planting material	-0.10**	-0.92**	0.42	0.34	401
	(-2.52)	(-2.11)			
Used certified planting material	-0.10**	-0.89*	0.40	0.32	401
	(-2.10)	(-1.79)			
Intercrop	-0.09	-0.78*	0.23	0.14	401
	(-1.56)	(-1.75)			
Times check crop	0.79	7.17	14.94	16.31	401
	(0.40)	(0.41)			
Weed in a timely manner	0.02	0.18	0.90	0.91	401
	(0.63)	(0.63)			
Remove and destroy infested material	0.19***	1.73***	0.32	0.48	401
	(5.75)	(3.01)			
Mulch	0.01	0.13	0.06	0.08	401
	(0.59)	(0.57)			
Apply ash	0.05**	0.43**	0.03	0.07	401
	(2.48)	(2.34)			
Spray with chillies	0.00	0.03	0.00	0.00	401
	(0.94)	(0.93)			
Prune	-0.10**	-0.88**	0.65	0.56	401
	(-2.23)	(-2.22)			
Change cycle	-0.05**	-0.46**	0.06	0.02	401
	(-2.18)	(-2.04)			
Burn crop residue pest/diseases control	0.00	0.02	0.01	0.02	401
	(0.29)	(0.30)			
Use trap crops	0.00	0.01	0.01	0.01	401
	(0.09)	(0.09)			
Grease	0.00	0.04	0.00	0.00	401
	(0.94)	(1.03)			

Note. Robust t statistics clustered at the clinic level in parentheses. All estimations control for lottery fixed effects. Impact estimates are 36-month impact estimates. Means are 36-month means. ITT = Intention to Treat. LATE=Local Average Treatment Effect, *p < .10; **p < .05; ***p < .01

Table A1.21 Impacts on Cultural Practices and Input Use: Coffee Producers (ANCOVA)

Dependent variable	ITT impact (1)	LATE impact (2)	Control mean (3)	Treatment mean (4)	N (5)
Organic fertilizer used	-0.04	-0.32	0.65	0.61	395
	(-0.92)	(-0.88)			
Inorganic fertilizer used	-0.04	-0.37	0.61	0.57	395
	(-0.77)	(-0.79)			
Pesticide used	-0.07	-0.60	0.57	0.53	395
	(-1.64)	(-1.40)			
Used certified planting material	-0.10**	-0.86	0.41	0.33	312
	(-2.28)				
Intercrop	-0.10	-0.82	0.24	0.14	312
	(-1.54)				
Times check crop	0.32	2.72	16.83	17.45	312
	(0.15)				
Weed in a timely manner	-0.03	-0.22	0.94	0.92	312
	(-0.72)				
Remove and destroy infested material	0.16***	1.36	0.31	0.46	312
	(3.01)				
Mulch	-0.02	-0.18	0.09	0.08	312
	(-0.67)				
Apply ash	0.03	0.28	0.05	0.08	312
	(1.33)				
Prune	-0.08	-0.68	0.68	0.59	312
	(-1.52)				
Change cycle	-0.04**	-0.36	0.06	0.01	312
	(-2.52)				
Burn crop residue pest/diseases control	0.01	0.05	0.01	0.02	312
	(0.57)	(0.60)			
Use trap crops	-0.00	-0.01	0.01	0.01	312
	(-0.13)				
Grease	0.00	0.04	0.00	0.00	312
	(0.93)				

Note. Robust t statistics clustered at the clinic level in parentheses. All estimations control for lottery fixed effects. Impact estimates are 36-month impact estimates. Means are 36-month means. ITT = Intention to Treat. LATE=Local Average Treatment Effect, *p < .10; **p < .05; ***p < .01

Table A1.22 Impacts on Production: Coffee Producers (Single Difference)

Dependent variable	ITT impact (1)	LATE impact (2)	Control mean (3)	Treatment mean (4)	N (5)
Value of production per acre	-0.10	-0.84	10.85	10.81	331
	(-0.70)	(-0.73)			
Cost of inorganic fertilizer per acre	-0.14	-1.14	9.23	9.04	207
	(-0.94)	(-1.05)			
Cost of pesticide per acre	0.03	0.17	9.38	9.34	186
	(0.11)	(0.15)			
Cost of labor per acre	-0.05	-0.28	8.81	8.75	175
	(-0.24)	(-0.26)			
Total costs per acre	0.02	0.23	10.36	10.53	226
	(0.14)	(0.15)			

Note. All outcome values are in natural logs. Robust *t* statistics clustered at the clinic level in parentheses. All estimations control for lottery fixed effects. Impact estimates are 36-month impact estimates. Means are 36-month means. ITT = Intention to Treat. LATE=Local Average Treatment Effect, *p < .10; **p < .05; ***p < .01.

Table A1.23 Impacts on Cultural Practices and Input Use: Kale Producers (Single Difference)

Dependent variable	ITT impact (1)	LATE impact (2)	Control mean (3)	Treatment mean (4)	N (5)
Organic fertilizer used	0.01	0.08	0.62	0.56	156
	(0.14)	(0.15)			
No. organic fertilizer applications	0.15	1.15	1.46	1.57	90
	(0.53)	(0.52)			
Inorganic fertilizer used	-0.09	-0.72	0.41	0.38	156
	(-1.29)	(-1.14)			
No. inorganic fertilizer applications	-0.01	-0.08	1.50	1.61	60
	(-0.02)	(-0.03)			
Family labor days (fertilizer)	-0.55	-3.64	2.24	2.02	85
	(-0.98)	(-1.31)			
Paid labor days (fertilizer)	-2.00**	-12.67	3.44	2.38	22
	(-2.29)	(-9.50)			
All labor days (fertilizer)	-0.54	-4.25	2.67	2.44	95
	(-0.93)	(-1.14)			
Pesticide used	-0.07	-0.61	0.41	0.40	156
	(-0.95)	(-0.92)			
No. pesticide applications	-1.44**	-9.15	3.28	2.43	62
	(-2.11)	(-1.35)			
Family labor days (pesticide)	-53.27	-324.3	51.05	2.67	51
	(-1.19)				
Paid labor days (pesticide)	-1.80	6.00	3.80	2.83	11
	(-1.15)	(0.97)			
All labor days (pesticide)	-44.61	-456.14	49.59	2.77	57
	(-1.22)	(-0.81)			
Crop rotation	0.13	0.95	0.31	0.41	133
	(1.10)	(1.34)			

Resistant Variety	-0.17**	-1.24**	0.52	0.41	133
	(-2.56)	(-2.39)			
Used improved planting material	-0.01	-0.05	0.67	0.63	133
	(-0.10)	(-0.11)			
Used certified planting material	-0.00	-0.00	0.78	0.67	133
	(-0.00)	(-0.00)			
Remove plant residue from prior harvest	0.05	0.36	0.87	0.84	133
	(1.52)	(1.48)			
Plant Early	0.05	0.39	0.52	0.59	133
	(0.53)	(0.57)			
Intercrop	0.05	0.37	0.07	0.10	133
	(0.99)	(1.04)			
Times check crop	-0.07	-0.51	14.78	12.53	133
	(-0.03)	(-0.04)			
Weed in a timely manner	-0.02	-0.12	0.89	0.87	133
	(-0.27)	(-0.29)			
Remove volunteer crops	-0.01	-0.10	0.70	0.68	133
	(-0.13)	(-0.15)			
Remove infested or damaged material	0.01	0.09	0.63	0.61	133
	(0.15)	(0.17)			
Mulch	-0.04	-0.32	0.20	0.22	133
	(-0.54)	(-0.58)			
Apply ash	0.01	0.07	0.13	0.14	133
	(0.23)	(0.26)			
Spray with chillies	0.03	0.21	0.00	0.05	133
	(1.60)	(1.62)			
Stake	-0.01	-0.09	0.02	0.01	133
	(-0.69)	(-0.72)			
Apply sand	-0.06	-0.40	0.06	0.01	133
	(-1.44)	(-1.55)			
Use trap crops	-0.04	-0.25	0.04	0.01	133
	(-1.31)	(-1.38)			
Burn crop residue pest/diseases control	0.01	0.09	0.06	0.04	133
	(0.24)	(0.27)			
Use traps	-0.01	-0.09	0.02	0.03	133
	(-0.42)	(-0.46)			
Cover with leaves	-0.04	-0.29	0.07	0.06	133
	(-0.53)	(-0.59)			

Note. Robust t statistics clustered at the clinic level in parentheses. All estimations control for lottery fixed effects. Impact estimates are 36-month impact estimates. Means are 36-month means. ITT = Intention to Treat. LATE=Local Average Treatment Effect, *p < .10; **p < .05; ***p < .01

Table A1.24 Impacts on Cultural Practices and Input Use: Kale Producers (ANCOVA)

Dependent variable	ITT impact (1)	LATE impact (2)	Control mean (3)	Treatment mean (4)	N (5)
Organic fertilizer used	0.01	0.10	0.62	0.56	152
	(0.16)	(0.18)			
No. organic fertilizer applications	0.14	1.07	1.46	1.57	90
	(0.47)	(0.47)			
Inorganic fertilizer used	-0.09	-0.73	0.41	0.38	152
	(-1.27)	(-1.14)			
No. inorganic fertilizer applications	0.02	0.30	1.50	1.61	60
	(0.06)	(0.07)			
Family labor days (fertilizer)	-0.57	-3.58	2.24	2.02	85
	(-1.00)	(-1.34)			
Paid labor days (fertilizer)	-2.01**	-8.68**	3.44	2.38	22
	(-2.25)	(-2.34)			
All labor days (fertilizer)	-0.55	-4.34	2.67	2.44	95
	(-0.95)	(-1.17)			
Pesticide used	-0.07	-0.58	0.41	0.40	152
	(-0.88)	(-0.87)			
No. pesticide applications	-1.43**	-8.93	3.28	2.43	62
	(-2.05)	(-1.36)			
Family labor days (pesticide)	-55.04	-336.58	51.05	2.67	51
	(-1.24)	(-0.98)			
Paid labor days (pesticide)	-1.80	6.00	3.80	2.83	11
	(-1.05)	(0.97)			
All labor days (pesticide)	-54.48	-571.37	49.59	2.77	57
	(-1.44)	(-0.79)			
Crop rotation	0.13	0.92	0.29	0.40	129
	(1.06)	(1.33)			
Resistant Variety	-0.18**	-1.31**	0.55	0.41	129
	(-2.54)	(-2.36)			
Used improved planting material	0.00	0.02	0.67	0.64	129
	(0.03)	(0.03)			
Used certified planting material	-0.02	-0.15	0.78	0.67	129
	(-0.41)	(-0.45)			
Remove plant residue from prior harvest	0.06*	0.38	0.86	0.85	129
	(1.70)	(1.59)			
Plant Early	0.04	0.28	0.53	0.60	129
	(0.40)	(0.43)			
Intercrop	0.05	0.35	0.08	0.09	129
	(0.94)	(1.01)			
Times check crop	-0.23	-1.59	14.27	12.65	129
	(-0.10)	(-0.12)			
Weed in a timely manner	0.01	0.06	0.88	0.87	129
	(0.14)	(0.15)			
Remove volunteer crops	-0.00	-0.03	0.71	0.68	129
	(-0.05)	(-0.05)			
Remove infested or damaged material	0.00	0.02	0.63	0.62	129

	(0.03)	(0.03)			
Mulch	-0.05	-0.30	0.20	0.22	129
	(-0.53)	(-0.57)			
Apply ash	0.01	0.07	0.14	0.14	129
	(0.23)	(0.26)			
Spray with chillies	0.03	0.21	0.00	0.05	129
	(1.60)	(1.63)			
Stake	0.01	0.05	0.00	0.01	129
	(1.09)	(1.26)			
Apply sand	-0.06	-0.42	0.06	0.01	129
	(-1.48)	(-1.59)			
Use trap crops	-0.04	-0.25	0.04	0.01	129
	(-1.30)	(-1.38)			
Burn crop residue pest/diseases control	0.02	0.11	0.06	0.04	129
	(0.32)	(0.36)			
Use traps	-0.01	-0.09	0.02	0.03	129
	(-0.42)	(-0.46)			
Cover with leaves	-0.04	-0.28	0.08	0.06	129
	(-0.54)	(-0.60)			

Note. Robust t statistics clustered at the clinic level in parentheses. All estimations control for lottery fixed effects. Impact estimates are 36-month impact estimates. Means are 36-month means. ITT = Intention to Treat. LATE=Local Average Treatment Effect, *p < .10; **p < .05; ***p < .01

Table A1.25 Impacts on Production: Kale Producers (Single Difference)

Dependent variable	ITT impact (1)	LATE impact (2)	Control mean (3)	Treatment mean (4)	N (5)
Value of production per acre	0.76	10.34	9.41	9.02	61
	(0.99)	(1.13)			
Cost of seed planted per acre	-1.31	-17.81	6.32	5.94	61
	(-1.37)	(-0.69)			
Cost of inorganic fertilizer per acre	-0.54	-7.28	4.10	4.21	61
	(-0.45)	(-0.45)			
Cost of pesticide per acre	-1.19	-16.21	4.45	4.62	61
	(-0.84)	(-0.60)			
Cost of labor per acre	1.68	22.85	2.24	2.90	61
	(1.31)	(0.66)			
Total costs per acre)	-0.91	-12.35	8.29	8.41	61
	(-1.12)	(-0.71)			
Gross margins per acre	1.43	53.05	1.87	-0.18	59
	(0.55)	(0.28)			

Note. All outcome values are in natural logs. Robust *t* statistics clustered at the clinic level in parentheses. All estimations control for lottery fixed effects. Impact estimates are 36-month impact estimates. Means are 36-month means. ITT = Intention to Treat. LATE=Local Average Treatment Effect, *p < .10; **p < .05; ***p < .01.

Table A1.26 Crop Production History: Maize Producers Only (Single Difference)

Dependent variable	ITT impact (1)	Control mean (2)	Treatment mean (3)	N (4)
Years HH has been farming	1.98 (0.26)	33.80	36.19	1,870
Years farming in location	1.35 (0.22)	30.21	32.24	1,872
Changed crops produced in last 5 years	0.01 (0.41)	0.20	0.20	1,797
Amount rainfall changed	0.00 (0.16)	0.87	0.87	1,797
Timing rainfall changed	-0.03** (-2.40)	0.96	0.93	1,797
Temperature changed	-0.03* (-1.91)	0.87	0.84	1,797
Number of insects increased	-0.02 (-0.80)	0.57	0.54	1,797
Number of diseases increased	-0.04 (-1.57)	0.57	0.53	1,797
Crop yields decreased	-0.04* (-1.77)	0.72	0.68	1,797
Pest information improved	0.04** (2.18)	0.32	0.37	1,797

Note. Robust *t* statistics clustered at the clinic level in parentheses. All estimations control for lottery fixed effects. Impact estimates are 36-month impact estimates. Means are 36-month means. ITT=Intention to Treat. *p < .10; **p < .05; ***p < .01.

Appendix 2. Mean Differences at Baseline for Attrition Analysis

In this appendix we investigate differential attrition at the 36-month follow-up by testing whether there is balance in terms of baseline characteristics for the sample of farmers who were still part of the sample at 36 months.

Table A2.1 Attrition Analysis for Roster Characteristic

	Control		Treatment		Balance Test		
Variables	Mean	N1	Mean	N2	Diff	SE	*p*-value
Household Size	4.11	1,694	4.20	856	0.09	0.17	0.62
% age between 0 and 5	0.10	1,694	0.10	856	0.01	0.01	0.35
% age between 6 and 12	0.14	1,694	0.15	856	0.01	0.01	0.27
% age between 13 and 18	0.12	1,694	0.13	856	0.00	0.01	0.88
% age between 19 and 35	0.29	1,694	0.27	856	-0.02	0.02	0.20
% age between 36 and 55	0.22	1,694	0.22	856	-0.00	0.01	0.83
% age between 56 and 70	0.08	1,694	0.08	856	0.00	0.01	1.00
% age 70 or older	0.04	1,694	0.04	856	0.00	0.01	0.69
Language used by respondent: English	0.05	1,694	0.03	856	-0.02**	0.01	0.03
Language used by respondent: Swahili	0.86	1,694	0.89	856	0.02	0.03	0.46
Language used by respondent: Other	0.09	1,694	0.09	856	-0.00	0.03	0.93
Household head gender	0.81	1,694	0.80	856	-0.01	0.02	0.79
Spouse gender	0.01	1,280	0.01	644	-0.00	0.00	0.84
Household head age	46.89	1,694	47.14	856	0.25	1.09	0.82
Spouse age	38.97	1,280	39.20	644	0.23	1.09	0.83
Household head attended school: Never attended	0.06	1,694	0.07	856	0.01	0.02	0.58
Household head attended school: Previously attended	0.93	1,694	0.93	856	-0.01	0.02	0.69
Household head attended school: Currently attending	0.00	1,694	0.00	856	-0.00	0.00	0.12
Spouse attended school: Never attended	0.06	1,694	0.06	856	0.01	0.02	0.72
Spouse attended school: Previously attended	0.70	1,694	0.69	856	-0.01	0.03	0.73
Spouse attended school: Currently attending	0.00	1,694	0.00	856	-0.00	0.00	0.69
Highest grade completed by household head: Std 8	0.20	1,694	0.21	856	0.01	0.02	0.51
Highest grade completed by household head: Form 6	0.02	1,694	0.02	856	-0.00	0.01	0.98
Highest grade completed by household head: University	0.01	1,694	0.01	856	0.00	0.00	0.60
Highest grade completed by spouse: Std 8	0.22	1,694	0.23	856	0.01	0.02	0.52
Highest grade completed by spouse: Form 6	0.00	1,694	0.00	856	0.00	0.00	0.31

Highest grade completed by spouse: University	0.00	1,694	0.00	856	-0.00	0.00	0.47
Household head can read	0.92	1,694	0.92	856	-0.00	0.02	0.98
Spouse can read	0.91	1,280	0.90	644	-0.01	0.02	0.66
Household head can write	0.87	1,694	0.85	856	-0.01	0.02	0.55
Spouse can write	0.84	1,280	0.80	644	-0.04	0.03	0.14

Note. Standard errors (SE) clustered at the plant clinic level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table A2.2 Attrition Analysis for Time Use and Labor

	Control		Treatment		Balance test		
Variables	Mean	N1	Mean	N2	Diff	SE	*p*-value
Household head earnings	143.85	1,649	54.61	825	-89.24	77.67	0.25
Spouse earnings	104.37	1,271	2.41	644	-101.96	63.90	0.11
Household head total hours worked	81.38	1,666	81.92	834	0.54	4.39	0.90
Spouse total hours worked	99.08	1,286	101.20	653	2.12	4.53	0.64
Household head hours spent on crop production	30.68	1,666	32.39	834	1.70	2.26	0.45
Spouse hours spent on crop production	29.74	1,286	30.84	653	1.10	2.23	0.62
Sum gross revenue from business	2,856.31	1,694	3,071.73	856	215.42	682.14	0.75
Sum net income from business	1,369.21	1,694	1,372.20	856	2.99	289.83	0.99

Note. Standard errors clustered at the plant clinic level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table A2.3 Attrition Analysis for Crop Production History

	Control		Treatment		Balance Test		
Variables	Mean	N1	Mean	N2	Diff	SE	*p*-value
Years HH has been farming	18.72	1,688	19.22	854	0.50	1.04	0.63
Years HH has been farming in current location	15.56	1,688	16.30	854	0.74	0.94	0.43
Produced different crops in the last 5 years	0.09	1,413	0.10	718	0.01	0.02	0.78
Amount rainfall changed	0.71	1,413	0.70	718	-0.01	0.04	0.75
Timing rainfall changed	0.83	1,413	0.85	718	0.02	0.03	0.55
Temperature changed	0.78	1,413	0.83	718	0.04	0.03	0.19
Number insects increased	0.43	1,413	0.46	718	0.02	0.05	0.66
Number diseases increased	0.40	1,413	0.44	718	0.04	0.05	0.46
Crop yields decreased	0.52	1,413	0.54	718	0.02	0.05	0.66
Pest information improved	0.13	1,413	0.10	718	-0.03	0.03	0.19

Note. Standard errors clustered at the plant clinic level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table A2.4 Attrition Analysis for Household Decision Making (Husband makes decision = 1)

Variables	Control Mean	N1	Treatment Mean	N2	Balance test Diff	SE	*p*-value
Making major household purchases?	0.54	1,694	0.44	856	-0.10**	0.04	0.02
Making purchases for daily household needs?	0.46	1,694	0.39	856	-0.06	0.05	0.17
How to use husband/male earnings off the farm?	0.51	1,694	0.43	856	-0.08*	0.04	0.06
How to use wife/female earnings off the farm?	0.44	1,694	0.38	856	-0.06	0.05	0.22
Which crops to produce for sale?	0.46	1,694	0.39	856	-0.07	0.05	0.15
Which crops to produce for home consumption?	0.44	1,694	0.38	856	-0.07	0.05	0.14
When to plan the crops?	0.45	1,694	0.38	856	-0.07	0.05	0.11
What inputs to buy for agricultural production?	0.46	1,694	0.39	856	-0.07*	0.04	0.10
When to apply fertilizers?	0.47	1,694	0.38	856	-0.09*	0.05	0.06
When to apply pesticides?	0.46	1,694	0.38	856	-0.08*	0.04	0.09
Who should work on crops being produced?	0.45	1,694	0.37	856	-0.08*	0.04	0.07
When to harvest?	0.42	1,694	0.35	856	-0.07	0.05	0.16
What to consume of production?	0.45	1,694	0.37	856	-0.08*	0.05	0.10
What to sell of production?	0.46	1,694	0.39	856	-0.07	0.05	0.12
Where to seek advice on crop production issues?	0.46	1,694	0.38	856	-0.08*	0.05	0.08
Where to seek advice on plant health issues?	0.47	1,694	0.39	856	-0.08*	0.05	0.09

Note. Standard errors clustered at the plant clinic level. Husband/male household member = 1. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table A2.5 Attrition Analysis for Pesticide Knowledge

Variables	Control Mean	N1	Treatment Mean	N2	Balance test Diff	SE	*p*-value
Consider preharvest interval is important	0.28	1,694	0.28	856	0.00	0.04	0.93
Regularly check for plant health problems	0.70	1,694	0.67	856	-0.03	0.03	0.35
Actions taken when plant health problems: None	0.29	1,694	0.33	856	0.04	0.04	0.34
Prefers chemical pest control	0.51	1,694	0.44	856	-0.07	0.05	0.15
Items used when spraying pesticides: Gloves	0.60	1,694	0.64	856	0.05	0.05	0.38
Items used when spraying pesticides: Mask	0.45	1,694	0.53	856	0.07	0.05	0.17
Items used when spraying pesticides: Goggles	0.19	1,694	0.21	856	0.02	0.03	0.45
Practices followed when spraying pesticides: Spray in morning	0.60	1,694	0.59	856	-0.01	0.04	0.88
Practices followed when spraying pesticides: Wash self afterwards	0.49	1,694	0.52	856	0.03	0.04	0.41
Practices followed when spraying pesticides: Use container	0.17	1,694	0.22	856	0.05*	0.03	0.05

Note. Standard errors clustered at the plant clinic level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table A2.6 Attrition Analysis for Crop Information Sources (Activity)

Variables	Control Mean	Control N1	Treatment Mean	Treatment N2	Balance test Diff	Balance test SE	Balance test *p*-value
Anyone in your family received advice on new seed varieties	0.14	1,694	0.11	856	-0.03	0.02	0.26
Anyone in your family received advice on pest control	0.13	1,694	0.11	856	-0.02	0.03	0.53
Anyone in your family received advice on fertilizer use	0.12	1,694	0.10	856	-0.02	0.02	0.30
Anyone in your family received advice on irrigation	0.03	1,694	0.02	856	-0.01	0.01	0.58
Anyone in your family received advice on marketing or crop sales	0.02	1,694	0.02	856	0.01	0.01	0.31
Anyone in your family received advice on postharvest technologies	0.01	1,694	0.01	856	-0.00	0.00	0.41
Anyone in your family received advice on value addition/agroprocessing	0.00	1,694	0.00	856	-0.00	0.00	0.12
New seed variety information useful	0.91	232	0.95	94	0.04	0.03	0.25
Pest control information useful	0.96	219	0.92	96	-0.05	0.04	0.25
Fertilizer use information useful	0.95	204	0.96	82	0.01	0.02	0.62
Irrigation information useful	0.91	44	0.88	17	-0.03	0.08	0.74
Marketing information useful	1.00	26	0.83	18	-0.17*	0.09	0.06
Postharvest tech information useful	0.83	18	1.00	6	0.17	0.12	0.18
Agroprocessing information useful	0.86	7	1.00	1	0.14	0.15	0.38

Note. Standard errors clustered at the plant clinic level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table A2.7 Attrition Analysis for Household Amenities and Conditions

Variables	Control Mean	Control N1	Treatment Mean	Treatment N2	Balance test Diff	Balance test SE	Balance test *p*-value
Roof made of: Iron sheets	0.92	1,694	0.93	856	0.01	0.02	0.62
Walls made of: Timber	0.24	1,694	0.28	856	0.04	0.06	0.50
Walls made of: Mud	0.20	1,694	0.20	856	-0.00	0.06	0.95
Walls made of: Concrete brick	0.11	1,694	0.09	856	-0.02	0.02	0.41
Floor made of: Mud/earth	0.52	1,694	0.48	856	-0.04	0.05	0.47
Floor made of: Concrete	0.41	1,694	0.45	856	0.04	0.06	0.47
Main source of drinking water: Directly from river/lake	0.43	1,694	0.36	856	-0.07	0.07	0.33
Main source of drinking water: Own tap	0.24	1,694	0.32	856	0.08	0.07	0.22
Distance from closest drinking source	1.40	1,694	1.34	856	-0.06	0.34	0.86
Household with electricity	0.23	1,694	0.26	856	0.02	0.05	0.66
Main type of energy used for cooking: Collected firewood	0.81	1,694	0.79	856	-0.01	0.04	0.73
Main type of toilet facility: Own pit latrine with slab	0.55	1,694	0.53	856	-0.02	0.04	0.52

Note. Standard errors clustered at the plant clinic level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table A2.8 Attrition Analysis for Credit Access

Variables	Control Mean	N1	Treatment Mean	N2	Balance test Diff	SE	*p*-value
Borrowed on credit from someone outside the HH	0.07	1,694	0.04	856	-0.03***	0.01	0.01
Debt from loans contracted in the past 12 months	53,346	127	27,841	35	-25,504*	12,872	0.05
Tried to borrow from someone outside the HH and turned down	0.01	1,694	0.01	856	-0.00	0.00	0.85
Would apply for a loan if certain he/she will get it	0.43	1,694	0.35	856	-0.08**	0.04	0.04

Note. Standard errors clustered at the plant clinic level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table A2.9 Attrition Analysis for External Support

Variables	Control Mean	N1	Treatment Mean	N2	Balance test Diff	SE	*p*-value
HH received money or goods from government programs	0.01	1,694	0.00	856	-0.00	0.00	0.23
Amount received from government programs	10,829	13	5,333	3	-5,495	3,585	0.15
HH received money or goods from NGO	0.00	1,694	0.00	856	0.00	0.00	0.50
Amount received from NGO	5,400	5	1,170	4	-4,230	2,684	0.15
HH received money or goods from individuals not from HH	0.00	1,694	0.00	856	-0.00	0.00	0.69
Amount received from individuals not from HH	7,666	3	180	1	-7,486	7,190	0.37
Total value of assistance	6,733	3	0.00	1	-6,733	7,659	0.44

Note. Standard errors clustered at the plant clinic level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table A2.10 Attrition Analysis for Social Capital

Variables	Control Mean	N1	Treatment Mean	N2	Balance test Diff	SE	*p*-value
Member of HH participated in: Farmers association	0.05	1,694	0.05	856	0.00	0.02	0.87
Member of HH participated in: Cooperatives	0.06	1,694	0.05	856	-0.01	0.01	0.60
Member of HH participated in: Common interest group	0.05	1,694	0.03	856	-0.01	0.01	0.21
Member of HH participated in: Water user association	0.02	1,694	0.02	856	-0.00	0.01	0.60
Member of HH participated in: Religious group	0.20	1,694	0.14	856	-0.06*	0.03	0.07
Member of HH participated in: Community welfare group	0.02	1,694	0.01	856	-0.00	0.01	0.59
Member of HH participated in: Social/investment group	0.01	1,694	0.01	856	0.00	0.00	0.87

Member of HH participated in: Cultural group	0.01	1,694	0.00	856	-0.01*	0.00	0.08
Member of HH participated in: Professional group	0.00	1,694	0.00	856	0.00	0.00	0.31
Member of HH participated in: Political organization	0.00	1,694	0.00	856	-0.00	0.00	0.80
How often participated in meetings: Farmers association	2.86	84	2.96	45	0.10	0.29	0.74
How often participated in meetings: Cooperatives	3.47	98	3.60	43	0.14	0.40	0.74
How often participated in meetings: Common interest group	2.13	78	2.50	28	0.37	0.30	0.22
How often participated in meetings: Water user association	2.64	33	2.85	13	0.21	0.53	0.70
How often participated in meetings: Religious group	1.16	342	1.24	119	0.08	0.09	0.35
How often participated in meetings: Community welfare group	2.66	29	2.00	12	-0.66	0.52	0.22
How often participated in meetings: Social/investment group	4.44	9	3.60	5	-0.84	1.26	0.51
How often participated in meetings: Cultural group	3.23	13	4.00	2	0.77	1.72	0.66
How often participated in meetings: Professional group	2.00	1	2.00	2	0.00		
How often participated in meetings: Political organization	3.00	7	3.33	3	0.33	1.46	0.83

Note. Standard errors clustered at the plant clinic level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Appendix 3. Plant Doctor Assessment – Regression Results

In this appendix, we formally estimate the impacts of PW-K training on plant health knowledge using linear regressions, which has some key advantages over the specification presented in Equations (1), (2), and (3). First, in addition to providing the exact same impact that one would obtain from those equations, linear regressions allow us to estimate the standard error of the impact, which is key to determine the statistical significance of the estimated effects. Second, using linear regressions allows us to control for other determinants of plant health knowledge, our outcome of interest. Including additional covariates not only allows us to estimate training impacts more precisely (i.e., to have lower standard errors), but it also lets us determine how sensitive the estimated impacts are to the inclusion of different sets of variables.

We estimate program impacts for the plant doctors trained in 2014 (G3) using the following econometric specification:

$$Score_{it} = \alpha + \beta_{11} d_{i,2015} + \beta_{12} d_{i,2017} + \beta_2 I_{i,G3} + \beta_{31} d_{i,2015} * I_{i,G3} + \beta_{32} d_{i,2017} * I_{i,G3} + \gamma X_{it} + \varepsilon_{it} \quad (A3.1)$$

where $Score_{it}$ is the total score obtained by person *i* in year *t*, with *t* being either 2014, 2015, or 2017; $d_{i,2015}$ and $d_{i,2017}$ are dummies equal to 1 if the observation is from 2015 or 2017, respectively, and 0 otherwise; $I_{i,G3}$ is a dummy equal to 1 if the assessed person was one of the new plant doctors in 2014 and 0 if the person is an untrained AEA (i.e., the person is part of the control group); X_{it} is a vector of covariates that includes information of the person assessed such as age, gender, education level, designation, deployment, specialization, experience, other non-PW-K trainings attended, and county fixed effects.

Note that the coefficients of interest are β_{31} and β_{32}, which measure the DD estimate for those trained in 2014 on scores in 2015 and 2017, respectively. Moreover, β_{11} is the difference in average scores between 2015 and 2014 for the control group ($\beta_{11} = \bar{Y}^{G1}_{2015} - \bar{Y}^{G1}_{2014}$), which measures to what extent untrained agents expand their knowledge on plant health issues over time. Analogously, β_{12} is the difference in average scores between 2017 and 2014 for the control group ($\beta_{12} = \bar{Y}^{G1}_{2017} - \bar{Y}^{G1}_{2014}$). In turn, β_2 measures the difference in average scores between the new plant doctors and the AEAs in 2014 ($\beta_2 = \bar{Y}^{G3}_{2014} - \bar{Y}^{G1}_{2014}$). This coefficient measures whether there were meaningful differences in plant health knowledge between G3 and the control group in 2014, before the new plant doctors had been trained. An estimate for β_2 close to zero means that new plant doctors and AEAs had similar levels of knowledge before the new plant doctors were trained.

Similarly, we estimate program impacts for the plant doctors trained in 2015 (G4) using the following regression:

$$Score_{it} = \delta + \theta_1 d_{i,2017} + \theta_2 I_{i,G4} + \theta_3 d_{i,2017} * I_{i,G4} + \gamma X_{it} + \zeta_{it} \quad (A3.2)$$

where as before $d_{i,2017}$ is a dummy equal to 1 if the observation is from 2017 and 0 otherwise; $I_{i,G4}$ is a dummy equal to 1 if the assessed person was one of the new plant doctors in 2015 and 0 if the person is an untrained AEA. Note that θ_3 measures the DD estimate for those PDs trained in 2015. θ_2 measures the difference in average scores between the 2015 new plant doctors and the AEAs in 2015 ($\theta_2 = \bar{Y}^{G4}_{2015} - \bar{Y}^{G1}_{2015}$). As before, this coefficient measures the difference in plant health knowledge between G4 and the control group in 2015, before the 2015 new plant doctors were trained. Again, an estimate for θ_2 close to zero means that new plant doctors and AEAs had similar levels of knowledge before the new plant doctors were trained.

Results

Table A3.1 presents the results of estimating the impacts of PW-K training on plant health knowledge (Equation A3.1) using the panel observations for the three groups of interest: the 2014 new plant doctors (G3), the current plant doctors (G2), and the AEAs (G1). The most important variables in the analysis are $d_{i,2015}$, $d_{i,2017}$, $I_{i,G3}$, and their interactions, $I_{i,G3} * d_{i,2015}$ and $I_{i,G4} * d_{i,2017}$. Column (1) estimates the differences in test scores between the three groups of agents in 2014, 2015, and 2017 without controlling for any observable characteristics. Columns (2) to (4) show how the estimated effects vary after controlling for different sets of observable characteristics. The regression in Column (2) controls for the counties where the officer works. In Column (3), we add a large set of observable characteristics, such as age, gender, designation, deployment, education, field of specialization, and years of experience as an extension officer, in addition to the county fixed effects. Lastly, in Column (4), we add two additional indices to the model estimated in Column (3) to account for non-PW-K trainings received in the last 12 months before the assessment and for consulting different information sources for plant health issues in the six months before the assessment.

The results in Column (1) for $d_{i,2015}$ show that the group of AEAs (control group or G1) scored only 0.36 fewer points in 2015 relative to their score in 2014, which means that the AEAs essentially had on average the same score in 2014 and 2015. Interestingly, the results for $d_{i,2017}$ show that the group of AEAs scored 9.2 more points in 2017 relative to their average score in 2014.[48] In turn, the results for $I_{i,G3}$ indicate that in 2014 the new plant doctors scored just 2.4 points above the AEAs and this estimate is not statistically significant. That is, in 2014, our baseline year, any untrained agent performed equally in the test, regardless of whether he or she had been selected to receive training later in that year. The estimated results for these variables remained almost unchanged after controlling for the different sets of covariates in Columns (2) to (4).

As discussed earlier, the 2015 DD estimate of the effect of trainings on the assessment for the doctors trained in 2014 (G3) is given by the interaction term between $d_{i,2015}$ and $I_{i,G3}$ and the 2017 DD estimate by the interaction between $d_{i,2017}$ and $I_{i,G3}$. The estimated results show that the plant doctors who were trained by PW-K in 2014 scored up to 7.4 more points in 2015 and 6.7 in 2017 relative to the AEAs.[49] Both of these estimates represent large and statistically significant differences in plant health knowledge. Indeed, given that the average AEA scored 51.9 points in the assessment (with a standard deviation of 10), scoring 7.4 additional points is equivalent to a 0.7-standard-deviation gain, which is considered a large effect in the assessment literature. Furthermore, note that the DD estimates are practically insensitive to the inclusion of the different observable characteristics in Columns (2) to (4). We interpret the estimated results for the 2014 new plant doctors as strong evidence that trainings provided by PW-K produce a large and significant effect on plant health knowledge.

Table A3.1 The Effects of PW Training on Plant Health Knowledge

Control Variables	(1)	(2)	(3)	(4)
d_{2015} (Year 2015)	-0.36 (0.28)	-0.38 (0.31)	-0.59 (0.54)	-0.60 (0.55)
d_{2017} (Year 2017)	9.22** (5.82)	9.23** (6.29)	9.23** (6.85)	9.28** (6.89)
I_{G3} (2014 New plant doctor)	2.39 (1.25)	2.43 (1.33)	1.44 (0.78)	1.33 (0.72)
d_{2015} * I_{G3} (same as DD_{2015}^{G3})	**6.92** (2.94)**	**6.94** (3.15)**	**7.34** (3.43)**	**7.35** (3.44)**
d_{2017} * I_{G3} (same as DD_{2017}^{G3})	**6.77* (2.52)**	**6.02* (2.40)**	**6.70** (2.73)**	**6.70** (2.73)**
I_{G2} (pre-2014 plant doctor)	9.28** (5.91)	8.22** (5.43)	8.36** (5.93)	8.44** (5.98)

[48] We conducted some additional analysis and establish that there was an overall increase in scores in the 2017 PDA. This increase in average scores was driven by the structured (open-ended) questions of the test and not the multiple-choice section. We conclude that more than an overall increase in plant health knowledge, the increase in average scores occurred because the professors who graded the assessment in 2017 were more generous with the grades given (i.e., the average score was higher for all participants compared to previous administrations of the assessment).
[49] These two DD estimates are not statistically different (p-value = 0.75). Thus, there is no evidence that the impacts of the training decreases over time.

Control Variables	(1)	(2)	(3)	(4)
$d_{2015} * I_{G2}$	-3.43	-3.50	-3.29	-3.28
	(1.69)	(1.76)	(1.83)	(1.82)
$d_{2017} * I_{G2}$	-3.12	-3.25	-3.79	-3.85
	(1.31)	(1.44)	(1.82)	(1.85)
Training Index (last 12m)				-0.11
				(0.27)
Consult Index (last 6m)				-0.43
				(1.12)
AEA mean score 2014	**51.94**	**51.94**	**51.94**	**51.94**
AEA mean score 2015	**51.58**	**51.58**	**51.38**	**51.38**
AEA mean score 2017	**61.17**	**61.17**	**61.31**	**61.31**
County fixed effects	**No**	**Yes**	**Yes**	**Yes**
Control variables	**No**	**No**	**Yes**	**Yes**
R^2	**0.27**	**0.33**	**0.44**	**0.44**
N	**606**	**603**	**587**	**587**

Note: t-stats reported in parentheses. Standard errors are robust to heteroskedasticity. Control variables included in Columns 2-4 include individual's age, gender, level of schooling, field of specialization, designation, years of experience as an extension agent, trainings received in the last year, use of extension methods, consultation activities with KARI, KEPHIS, NGOs, colleagues, and agrodealers, and deployment. Omitted categories are PD – AEA; Designation – Agricultural Assistant (JAA, AA I, SAA, CAA); Deployment – FEO; Education – Diploma; Specialisation – Agriculture. * p<0.1; ** p<0.05; *** p<0.01

We also estimate the effects of PW-K trainings for those plant doctors who at the time of first PDA in 2014 had already been trained by PW-K. To do so, we expand Equation A3.1 and include the variable $I_{i,G2}$, a dummy variable that takes the value of 1 if the extension officer was a current plant doctor already in 2014 and 0 for the other type of extension officers (i.e., AEAs and new plant doctors). We also include interactions of $I_{i,G2}$ with the year dummies to measure how the average scores for the current plant doctors (relative to the AEAs) change over time. The results for $I_{i,G2}$ show that in 2014, the then current plant doctors scored between 8 to 9 points higher than AEAs. Although we don't have assessment data for these two groups before the current plant doctors were trained the first time (i.e., we don't have baseline data that allows us to implement a DD strategy for G2), it is encouraging that these estimates are similar to the effect of training that we find for the 2014 new plant doctors in 2015 and in 2017. That is, whereas the estimated effect of the program in Column (4) for G2 is 8.44 points in 2014, the DD estimate of the 2014 new plant doctors (G3) in 2015 is 7.35 points, a difference that is not statistically significant. Lastly, the results in Column (4) for the interaction of being a current plant doctor (G2) and the 2015 dummy reveals that current plant doctors outperformed AEAs by approximately 5 test points (=8.44-3.28) in 2015 and in 2017 (=8.44-3.85), which is equivalent to 0.5 standard deviations, again a significant difference.

We also estimated equation A3.2 to determine the impact of PW training on plant health knowledge for those plant doctors who were trained in 2015 for the first time (G4). The results are strikingly similar to the DD estimates for the cohort of plant doctors trained in 2014 (G3). Indeed, the results show that in 2015 (the baseline for this analysis) officers who had been selected to be plant doctors scored only 0.36 points (t-stat = 0.23) higher than the control group (G1). That means that, as expected, the treatment and control groups at baseline had the same level of plant health knowledge. Moreover, the DD estimate shows that officers trained in 2015 (G4) scored 7.05 points (t-stat = 2.88) higher than the control group. This estimated impact is very similar and not statistically different from the DD estimated reported in Table A3.1 for the plant doctors trained in 2014 (G3).

Overall, we interpret the similarity of the estimated training impacts on plant health knowledge across years and different cohorts of plant doctors (G3 and G4) as strong as evidence that trainings provided by PW-K produce a large and significant effect on plant health knowledge.

Appendix 4. Local Average Treatment Effect (LATE) Impact Estimates

In any experimental evaluation, across both the treatment and control groups, there are three types of individuals (1) those who always receive the program regardless of treatment status, (2) those who never receive the program regardless of treatment status, and (3) the compliers, those who only receive the program because of having been assigned to the treatment group. In examining the treatment effect on the treated, we would ideally examine the effect on the compliers, those who receive the program because of their assignment into the treatment group. However, because it is not possible to observe all the types of individuals, we must rely on the additional estimation methodologies.[50] The local average treatment effect (LATE) measures the impact of the treatment on this third subset of people, the compliers, whose treatment status changes as a result of being assigned to the treatment.

The LATE estimate scales the intention-to-treat estimate by one over the fraction of the sample who are compliers – that is those who were affected by the assignment. We will use econometric methods to estimate what that fraction is, which we describe in more detail below. To see how the LATE estimate corresponds to the intention-to-treat estimate, consider the following algebraic relationships. Define the share of those who were always treated as $\emptyset_A$ and let Y_A^1 represent the outcome for the always-treated when they attend the plant clinic. Define the share of those who had never been treated as $\emptyset_N$ and let Y_N^0 represent the outcome for the never-treated when they do attend the plant clinic. Finally, define the share of

[50] Specifically, the first group of farmers who always receive the program includes farmers who, despite being in the control group, attend the plant clinic as well as farmers in the treatment group who attend the clinic but would have done so regardless of their treatment status (i.e., even if they had been assigned to the control group). While we can observe those control group farmers who attend the plant clinics, we are not able to observe which of the treatment farmers would have attended had they been in the control group. The second group of farmers who never receive the program includes farmers who, despite being in the treatment group, never attend the plant clinic as well as farmers in the control group who never attend the clinic but would not have attended the clinic regardless of their treatment status (i.e., even if they had been assigned to the treatment group). While we can observe those treatment farmers who do not attend the plant clinics, we are not able to observe which of the control farmers would not have attended had they been in the treatment group. We are also unable to observe the third group of farmers, the compliers, that includes treatment group farmers who attend the plant clinic but would not have attended if they had been assigned to the control group as well as the control group farmers who do not attend the plant clinic but would have if they had been assigned to the treatment group.

compliers as $\emptyset_C$ = (1 - $\emptyset_N$ - $\emptyset_A$); let Y_C^0 represent the outcome for the compliers in the control group who do not attend the plant clinics and Y_C^1 represent the outcome for the compliers in the treatment group who do attend. With these definitions, we can construct the following equations:

The mean outcome for those assigned to the control group is

$$Y^c = \emptyset_A Y_A^1 + \emptyset_C Y_C^0 + \emptyset_N Y_N^0 \quad \text{(A4.1)}$$

The mean outcome for those assigned to the treatment group is

$$Y^T = \emptyset_A Y_A^1 + \emptyset_C Y_C^1 + \emptyset_N Y_N^0 \quad \text{(A4.2)}$$

The difference between these is

$$Y^T - Y^c = \emptyset_C * (Y_C^1 - Y_C^0) \quad \text{(A4.3)}$$

Thus, the LATE estimate, the impact of the treatment on the compliers, is equal to

$$\frac{(Y^T - Y^c)}{\emptyset_C} = Y_C^1 - Y_C^0 \quad \text{(A4.4)}$$

That is, the LATE estimate is equal to the intention-to-treat estimate divided by the fraction of the sample who are affected by the assignment. Because the LATE method estimates the impact of CABI for those farmers in the treatment group whose assignment into the treatment induced them to attend the clinic, the effect size is larger than the intent-to-treat estimates. In the outcome regression tables, we calculate the LATE using an instrumental variables approach in which we use the treatment assignment as an instrument to predict whether the farmer attended the plant clinic. Using the treatment assignment to predict clinic attendance, enables us to obtain the fraction of compliers. However, it is worth nothing that the additional estimation required in this approach increases the standard error of the estimates; so although the LATE impact estimates are larger than the intent-to-treat estimates, sometimes they may be statistically indistinguishable from zero.

While the above equations document the concept of the LATE estimate, in practice the LATE estimate is often calculated by an instrumental variables or two-stage least squares approach where, in the first stage, we estimate for treatment receipt (in our case, clinic attendance) based on the randomized treatment assignment. As explained above, this first stage enables us to estimate the fraction of the population that are compliers. The first stage enables us to generate predicted values for actual treatment receipt, where the caret symbol indicates the value is estimated. We can then use the predicted values of treatment receipt in our estimate of the local average treatment effect in a regression, using the treatment receipt instead of randomized assignment. Note that actually conducting the two-stage procedure would produce incorrect standard errors. Statistical packages like Stata have built-in modules like the "Burn crop residue for pest/disease control" ivregress, which automatically correct the standard errors. By using the randomized assignment as an instrument to predict clinic attendance, rather than using the observed clinic attendance of all, we are able to examine the impact for those who attended because the experiment induced them to attend.

The results of the first-stage regression of clinic attendance on treatment status are presented in Table A5.1. These results show that being in the treatment group in 2017 strongly predicts clinic attendance. In fact, farmers randomized to the treatment were 8 percentage points more likely to attend the plant clinic.

Table A4.1 First Stage Regression

	Clinic attendance
Treatment status	0.08***
	(5.70)
Mean treatment	0.14
Mean control	0.06
F-stat (treatment)	32.52
R^2	0.06
N	2550

Note. Robust *t* statistics clustered at the clinic level in parentheses. Estimation controls for lottery fixed effects. *p < .10. **p < .05. ***p < .01.

To estimate a LATE without bias, some identification assumptions need to hold (Duflo, Glennerster, and Kremer, 2008). First, the comparison of outcomes for farmers in both treatment and control groups identify the causal impact of the instrument. This is true by construction in the evaluation of PW since the instrument (i.e., being assigned to a treatment catchment area) is randomly assigned. The second assumption, known as the exclusion restriction, is that potential outcomes are not *directly* affected by the instrument.[51] This assumption, however, does not necessarily hold in all randomized evaluations and we think this is the case in the evaluation of PW. The exclusion restriction assumption implies that the instrument does not affect the outcomes of those farmers who, despite living in treatment areas, decide not to attend plant clinics. However, one fundamental characteristic of PW is that trained plant doctors interact with farmers, other extension agents, and supervisors not only thought plant clinics, but also outside the plant clinics. They transmit knowledge that reflects their new training as plant doctors and what they learn from farmers in the clinics. These interactions also influence the activities of local

[51] There is a third assumption called the monotonicity assumption. It states that the instrument needs to make *every* farmer either weakly more (or less) likely to actually participate in the program. That is, every farmer in the treatment group is no less likely to attend a plant clinic than had they been in the control group. While this assumption is untestable, in most cases it is reasonable to assume it holds.

field agricultural extension offices. For example, agricultural extension offices hold farmer field days to promote certain agricultural practices, the focus of which could change based on new information from plant doctors. Moreover, farmers in treatment areas who do not attend plant clinics can also benefit from the fact that their crops may be healthier if neighbors who attend plant clinics have healthier crops as a result of clinic attendance. Thus, the fact that PW activities may directly affect not only plant clinic users, but more importantly non-users of plant clinics in treatment areas imply that the estimation of a LATE can be severely (upward) biased. While the violation of the second assumption cannot be tested empirically, its implications need to be carefully considered to avoid reintroducing biases in the estimation of program impacts.

AMERICAN INSTITUTES FOR RESEARCH®

Established in 1946, American Institutes for Research (AIR) is an independent, nonpartisan, not-for-profit organization that conducts behavioral and social science research on important social issues and delivers technical assistance, both domestically and internationally, in the areas of education, health, and workforce productivity.

MAKING RESEARCH **RELEVANT**

AMERICAN INSTITUTES FOR RESEARCH
1000 Thomas Jefferson Street NW
Washington, DC 20007-3835 | 202.403.5000
www.air.org

LOCATIONS

Domestic: Washington, DC (HQ) | Monterey, Sacramento, and San Mateo, CA | Atlanta, GA | Honolulu, HI | Chicago and Naperville, IL Indianapolis, IN | Metairie, LA | Waltham, MA | Frederick and Rockville, MD | Chapel Hill, NC | New York, NY | Columbus, OH | Cayce, SC Austin, TX | Reston, VA

International: El Salvador | Ethiopia | Haiti | Honduras | Kyrgyzstan | Tajikistan | Zambia

3869_03/18