Independent assessment of TechnoServe's coffee agronomy training program in Rwanda

Final report

FEBRUARY 2013

laterite
DATA | RESEARCH | ADVISORY

Table of Contents

Chapter 1. Introduction 7
Chapter 2. Motivations and Objectives of study 9
Chapter 3. Methodology & Overview of Findings 14
Chapter 4. Yield Impact Results 18
Chapter 5. "Best Practice" Adoption Results 35
Chapter 6: The Monitoring Effect 57
Chapter 7: Conclusion & Recommendations 64
Annex A: Focus Group & Semi-Structured Interview Transcripts 68
Annex B: Technical Annexes 78

Tables

Table 1: Total number of coffee farmers participating in the Rwanda Agronomy Program as of October 2012 8
Table 2: Structure of TechnoServe Coffee data 9
Table 3: Structure of Yield and BP data 10
Table 4: Average yields (kg/tree) in program Cohorts 10
Table 5: Estimated cumulative impact of training on average yields (kg/tree) 11
Table 6: Best Practice Indicators 13
Table 7: Baseline Yields in 2010, 2011 and 2012 Cooperatives 19
Table 8: Measuring the similarity of the cohorts 22
Table 9: Average Yields data for 2010, 2011 and 2012 Cohorts in Sub-Sample 27
Table 10: Regression results estimating impact of time 28
Table 11: Impact Estimates 2010 31
Table 12: Before- and after- training yields estimates in sub-sample 32
Table 13: Reported Best Practice Adoption Data 36
Table 14: Best Practice groups of prior knowledge 38
Table 15: Structure of Best Practice Data 41
Table 16: Measuring the similarity of "high-attendance" vs. "low-attendance" farmers 42
Table 17: Population vs. Best Practice Sample Characteristics 43
Table 18: Potential Over-estimation Scenarios 43
Table 19: Difference between reported FT data and Enumerator data 44
Table 20: Spot Check Analysis Sample 45
Table 21: Reported vs. Spot-Check Data 45
Table 22: Observed Non-Adoption vs. Adoption rates (1) 47
Table 23: Observed Non-Adoption vs. Adoption rates (2) 47
Table 24: BP Adoption rates (Yield vs. Non-Yield Sample) 60
Table 25: Suggested changes to Best Practices Adoption metrics 66

Figures

Figure 1: Cumulative distribution of yields in baseline and treatment samples 21
Figure 2: Yield levels in one year compared to yield levels in the previous year 24
Figure 3: Yield levels in one year compared to yield levels two years earlier (Cohort 2010, t=2012) 25
Figure 4: Cumulative distribution of baseline yields (sub-sample limited to yield>0.6 and yield<3.5) 26
Figure 5: Distribution of yield data before and after one year of training 32
Figure 6: Relation between BPSim and differences in yield improvements between pairs of farmers (Cohort 2010)....33
Figure 7: Link between prior knowledge and best practice adoption (C2010) 38
Figure 8: Best practices with seasonal behaviour (Cohort 2010) 40
Figure 9: Best practices with increasing trends (Cohort 2010) 40
Figure 10: Best practices with increasing trends (Cohort 2010) 40
Figure 11: Predicted errors by Best Practice 48
Figure 12: Average attendance rate by number of best practices adopted (Cohort 2010) 50
Figure 13: Likelihood of adopting best practices based on attendance levels 50
Figure 14: Attendance rate of adopters vs. non-adopters by farmers session (Cohort 2010) …………………………..51
Figure 15: Difference in adoption rates between trained and untrained farmers by session (Cohort 2010) 51
Figure 16: Estimated marginal increase in the likelihood of adopting best practices based on attendance levels 52
Figure 17: Difference in adoption rates of trained and untrained farmers (Cohort 2010) 53
Figure 18: Difference in Nutrition adoption rates between trained & untrained farmers by session (Cohort 2010) 54
Figure 19: Matrix of most effective training sessions based on theory vs. in practice 54
Figure 20: Attendance rates in yield sample vs. comparison group and placebo 58
Figure 21: Attendance rates in yield samples vs. comparison group (Cohorts 2010 & 2011) 59
Figure 22: Attendance rates in BP sample vs. comparison group (Cohorts 2010) 61

Executive Summary

This report presents an independent impact assessment of TechnoServe's Agronomy program in Rwanda and provides compelling evidence that the Agronomy program has had a substantial impact on both best practice adoption and the productivity of coffee trees in participating cooperatives.

The Rwanda Coffee Agronomy Program was designed to help farmers increase the productivity of their coffee trees by building their knowledge and skills in sustainable and yield-increasing agricultural practices. The project provides participants with a structured 2-year training program, which runs monthly in the first year of training and twice per month in the second year. To date more than 20,000 farmers have completed the training - in batches of about 7,500 farmers each year (called Cohorts) - with another 8,500 currently in training. The training curriculum is structured around known sustainable coffee-farming practices that improve the productivity of coffee trees and reduce their cyclicality. The eleven best practices that the program monitors and focuses on can be grouped into four inter-related categories:

1. ***Maintaining the plot***, through mulching, weeding, and ensuring there is sufficient shade for the coffee trees;
2. ***Caring for the coffee trees***, by pruning them regularly and rejuvenating every 6-7 years;
3. ***Providing the right inputs mix***, in particular through compositing and better nutritional practices; and lastly,
4. ***Using sustainable farming methods***, by limiting soil erosion, making safe use of pesticides and finally keeping detailed records to better manage farming activities.

Farmer from Kinyaga Cooperative
"Long time ago, the agronomy government official used to come and just tell us to do this and that but TechnoServe came all the way to the field and showed us how to actually do it. Now we know how to use NPK, we know how and when to pick our berries from the coffee tree and now we get more yields from the same tree."

Farmer from Cocamu Cooperative
"We used a lot of traditional techniques that we had learnt from our parents and friends but it's only since TechnoServe came that we received professional training."

The objective of the study is to (i) test the robustness of current impact estimates; (ii) provide new insights on yield impact and best practice adoption; and (iii) independently verify data collection methods with the objective of identifying inherent biases in TechnoServe's current M&E approach. Four complementary sets of tools are used throughout:

- Internal validity checks to identify potential biases within the identification strategy;
- Field spot-checks to test the accuracy of reported best practice adoption data;
- Specific strategies to overcome identified biases and answer the question of whether the program had an impact on coffee tree yields and best-practice adoption in participating cooperatives; and
- Semi-structured interviews and focus groups with coffee farmers to test some of the initial findings in the field.

Impact of the of program on coffee yields

We follow a three step process to determine whether there is sufficient evidence to support the claim that the Agronomy program has had an impact on average yield levels: (i) we first study potential biases (see below); (ii) second, we propose a strategy to overcome some of the identified sources of bias and provide an estimate of the impact of the coffee program on a specific sub-sample of farmers; and (iii) lastly, seek to establish a link between yield increases and best practice adoption.

We find strong evidence of a positive impact of the training program on coffee yields. For a sub-group of farmers (that are representative of more than 70% of the training population), **we estimate that the Agronomy training was associated with a 72.6% increase in average yield levels for farmers in Cohort 2010 after two years of training**. Initial results suggest that the impact of the training program on Cohort 2011 could be even larger: after one year of training yield levels in Cohort 2011 had increased by 75.5% compared to 57.5% in Cohort 2010, even though these initial estimates are biased upwards by the annual cyclicality of coffee trees. An innovative technique based on the similarity of best practice adoption patterns between pairs of farmers also shows that yield increases are intricately related to best practice adoption. While the results are clearly indicative of a causal impact, it is important to note that these impact estimates might be over-stating the actual impact of the program, as it is not possible to distinguish between the effect of the training program and the additional effect of monitoring (an unexpected finding of this study - see below).

Impact of the program on best practice adoption

Given the structure of the data and in particular the lack of a baseline, we are not able to provide a formal estimate of the impact of the program on best practice adoption. Instead, we build a compelling evidence-based case as to why the program is almost certain to have had an impact on best practice adoption. To do this we follow a three step process: (i) first, we check for selection bias; (ii) second, we check the validity of reported best practice adoption rates - testing for Farmer Trainer over-reporting, in particular through field spot checks on a randomly selected group of farmers in "high risk" cooperatives; and (iii) lastly, build the case for the link between training and best practice adoption, by comparing attendance rates and trends to best practice adoption patterns.

We find a strong association between attendance and best practice adoption in all Cohorts, which is further evidence of the potential impact of the training program on best practice adoption. **The higher a farmer's attendance rate, the more likely he/she will adopt a best practice.** Supporting the hypothesis that the training program leads to higher adoption, we find that as the training progresses the difference in adoption rates between "trained" farmers and "untrained" farmers increases. Moreover, there appears to be a clear link between attending a specific training session on a certain best practice, and adopting the corresponding best practice. These quantitative findings are supported by anecdotal evidence from the field: for all best practices, more than 50% of farmers interviewed as part of a random spot-check claimed have acquired these best practices through TechnoServe's training program.

Verification of data collection methods and biases

The design of TechnoServe's M&E strategy is based on the premise that old and new cohorts should resemble each other on average and that, therefore, subsequent program Cohorts of farmers (that have not yet been trained) can be used as control regions for current Cohorts (that are already undergoing the training). We identify five key risks and sources of potential bias in TechnoServe's M&E system:

1. ***The risk of selection bias due to the way successive Cohorts and M&E samples were constructed***. We find selection bias at three levels: (i) there are small but significant differences between Cohorts on both farmer and cooperative level indicators; (ii) the nature of cooperatives that make up Cohort 2009 is different from Cohorts 2010, 2011 and 2012, leading to Cohort 2009 being dropped from the analysis; and (iii) the best practice sample is not representative of the training population, but rather a sub-sample of farmers that attended 50% of sessions in year 1. We tackle selection bias by controlling for farmer, cooperative, sector and district-level indicators, including both socio-economic and agro-climatic controls, and by working with a sub-sample of farmers.

2. ***Omitted observations in time***. The fact that we do not have yield baseline data for the control groups makes it impossible to conduct a typical difference-in-difference impact analysis and to check the

internal validity of the evaluation by comparing baseline data in the control and treatment groups. To circumvent this problem, we exploit the structure of TechnoServe's yield data to estimate the effect of time and adjust the results accordingly.

3. ***Over-reporting by Farmer Trainers, who conducted both the training and the data collection***. We do not find sufficient evidence to support the claim that there has been systematic over-reporting of best practice adoption or yield levels by Farmer Trainers. A simple comparison of adoption rates reported by Farmer Trainers and TechnoServe enumerators, suggests there are no significant differences between them. Moreover, while spot-checks on a sample of 270 randomly selected farmers suggest that Farmer Trainers are over-reporting results by 20 percentage points, a deeper analysis of the data shows that the most likely explanation to this difference is not Farmer Trainer over-reporting but instead enumerator error.

4. ***Indicator design, leading to an overestimation of actual adoption.*** TechnoServe collects data on 11 best practices. For certain indicators we show that the way results are reported can lead to an over-estimation of actual adoption rates.

5. ***The effect of monitoring on the way famers experience the training program***. One of the most unexpected and surprising results of this study is the large effect of monitoring on the way farmers experience the training program. TechnoServe's extensive and unique monitoring and evaluation system in Rwanda enables us experimentally conduct one test that even the best randomized control trials would not allow: to test the impact of the monitoring and evaluation system itself on project beneficiaries. We find that the regular monitoring of farmers in the yield sample – which consists of bi-monthly (twice per month) interactions between the farmer and the project staff for three years – led to a 12 to 15 percentage point increase on farmer attendance rates, a 7 percentage point increase in best practice adoption and a significant increase in fertilizer usage. Similarly, monitoring on farmers in the best practice sample - which only happens twice per year - led to a temporary 5 percentage point increase in their attendance rates. These results lead to interesting questions on how to best leverage the monitoring effect to increase program outputs. Creating mechanisms to provide the 'illusion' of monitoring should be piloted and tested as they might increase farmers' attendance and adoption rates.

Recommendations relating to M&E framework

We do not recommend any major changes to TechnoServe's monitoring and evaluation system, but do recommend adjustments. Proposed adjustments include altering the sampling frame for the best practice sample to include all farmers, not just high attendance farmers; adjusting the timing of best practice data collection, in particular by capturing baseline information on best practice adoption using comparable metrics; and finally ensuring that the best practice metrics are more specific. The objective of these recommendations are to increase the comparability of the data and the variation within the samples – this will enable TechnoServe to conduct better measure impact and use its monitoring and evaluation system more effectively to improve program design.

Chapter 1. Introduction

This report presents an independent impact assessment of TechnoServe's Agronomy program in Rwanda, focusing in particular on the impact of the program on coffee yields in participating cooperatives and on the adoption of best-practice farming techniques.

TechnoServe's East Africa Coffee Initiative, operating in Rwanda, Kenya, Tanzania, and Ethiopia, was started in 2008 with funding from the Bill and Melinda Gates foundation. The project was designed to increase the incomes of smallholder coffee farmers and enable them to participate in the specialty coffee value chain, through two interventions: a Wet-mill program and an Agronomy program. Phase I of the project, called "Doubling Coffee Incomes for 1 million East African Smallholder Farmers Project", came to an end in December 2011. The second phase of the project, called the "Increasing Coffee Incomes for East African Smallholder Farmers Project", is currently ongoing and targets an expansion of project activities in Ethiopia and a consolidation of results in Rwanda, Kenya and Tanzania. We focus here on the results of Phase I of the Agronomy program in Rwanda and ongoing activities through to October 2012.

The objective of TechnoServe's Agronomy Program is to help farmers increase the productivity of their coffee trees by building their knowledge and skills in sustainable and yield-increasing agricultural practices through a two-year training program. Training is delivered following a very structured and hands-on approach. The structure of the program can be summarized as follows:

- Each year TechnoServe selects a handful of cooperatives to participate in the program, based on a set of criteria (specified in subsequent sections). This group of cooperatives is called a "Cohort".
- There are about 1000 registered participants in each cooperative location, with farmers self-selecting into the program. Most participants are members of the cooperative but the program is also open to non-members.
- Based on their location, farmers in each cooperative are divided into groups of about 30 people (the exact size depends on demand, the number of farmer trainers, and has evolved due to financial considerations since the project started). Each group selects a respected coffee farmer in the community to represent them; in the terminology of the project, these representatives are called "Focal Farmers".
- Hands-on training to each group of farmers is provided by a TechnoServe "Farmer Trainer". The training takes place in the coffee field of the "Focal Farmer" once per month in the first year of training and every two months in the second year. The responsibility of the "Focal Farmer" is to ensure that participation in his/her group of farmers is high and to make a part of his/her plot available for the training.
- In addition to training about 300-500 farmers each, TechnoServe's "Farmer Trainers" maintain demonstration plots in each cooperative to showcase and test the impact of best farming practices.

The training curriculum is structured around known sustainable coffee-farming practices that improve the productivity of coffee trees and reduce their cyclicality. The eleven best-practices that the Agronomy program monitors and focuses on can be grouped into four inter-related categories:

1. ***Maintaining the plot***, through mulching, weeding, and ensuring there is sufficient shade for the coffee trees;
2. ***Taking care of the coffee trees***, by pruning them regularly and rejuvenating every 6-7 years;
3. ***Providing the right inputs mix***, in particular through compositing and better nutritional practices; and lastly,
4. ***Using sustainable farming methods***, by limiting soil erosion, making safe use of pesticides and finally keeping detailed records to better manage farming activities.

By October 2012, four years after the start of the program, over 11,000 farmers had completed the program and 18,000 farmers were either in year 1 or 2 of the program (see table 1). This amounts to about 10% of Rwanda's active coffee farming population.[1] During this period TechnoServe had either worked with or helped create 28 coffee cooperatives in all coffee producing areas of the country (see map). By design, none of the Cohorts were geographically concentrated; i.e. each Cohort included cooperatives in all provinces of the country, excluding Kigali.

TechnoServe's Agronomy Clients (by October 2012)

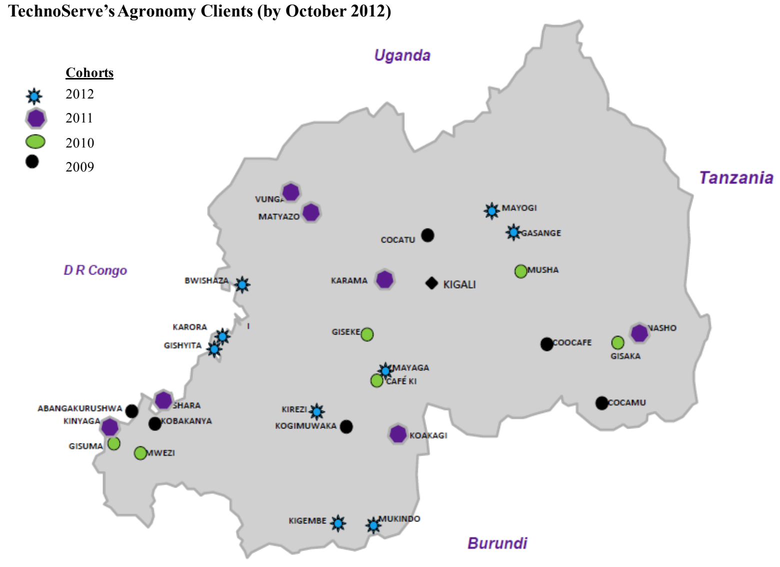


Table 1: Total number of coffee farmers participating in the Rwanda Agronomy Program as of October 2012

Cohort	Status	Registered Farmers	Farmers attending regularly
2009 Cohort	Training Completed	4,147	3,433
2010 Cohort	Training Completed	9,123	7,756
2011 Cohort	Year 2 of Training	9,652	8,274
2012 Cohort	Year 1 of Training	8,500	TBD

This study provides compelling evidence that the Agronomy program has had substantial impact on both best practice adoption and the productivity of coffee trees in participating cooperatives, thereby largely achieving project objectives. Before delving into the details of the analysis, we briefly discuss the motivation and objectives of this work and the methodology adopted.

[1] Estimate based on 2009 Coffee Census results, which found that there were 394,000 coffee farmers in Rwanda.

Chapter 2. Motivations and Objectives of study

The main objective of the study was to provide an independent assessment of the impact TechnoServe's Agronomy program in Rwanda on the productivity of coffee trees and the adoption of sustainable coffee-farming practices in participating cooperatives, building on TechnoServe's rich project database, backed-up with independent data collection. In particular, this study was designed to: (i) test and check the robustness of current impact estimates; (ii) to provide new insights on yield impact and best practice adoption; and (iii) to independently verify data collection methods and identify inherent biases in TechnoServe's current M&E approach.

a. TechnoServe's M&E system

Over the past few years, TechnoServe has put in place an elaborate system to monitor the performance of its Agronomy training program, focusing on three key variables: (i) attendance rates, (ii) best-practice adoption rates and (iii) coffee tree productivity.

(i) Attendance Data: Individual farmer attendance data is collected during every training session by the "Farmer Trainer", making it possible to identify exactly which training sessions a specific farmer attended and which he/she did not.

(ii) Best-Practice Adoption Data: Information on best-practice adoption is collected twice per year from a randomly selected sub-sample of farmers that attended at least 50% of classes (sample sizes range from 500 to 1000 farmers, depending on the Cohort).

(iii) Yield Data: Data on the productivity of coffee trees along with best practice adoption data is collected on a monthly basis from a smaller randomly selected group of farmers (300-500 farmers per Cohort).

Combined, we will show that these datasets provide sufficient information to conduct an impact assessment of TechnoServe's Agronomy program in Rwanda (see table 2).

Table 2: Structure of TechnoServe Coffee data

Data	Cohort	Year Collected	Sample (# of farmers)
Attendance data	2009 Cohort	2009-2010	4173
	2010 Cohort	2010-2011	9124
	2011 Cohort	2011-2012	9697
	2012 Cohort	2012-2013	8500
BP Adoption data	2009 Cohort	Round 1 2011	292
	2010 Cohort	Round 1 2011	968
	2009 Cohort	Round 2 2011	489
	2010 Cohort	Round 2 2011	1052
	2011 Cohort	Round 1 2012	731
	2012 Cohort	Round 1 2012	435
Yield data	2009/10 Cohort	2010	381
	2009/10 Cohort	2011	542
	2011 Cohort	2011	351
	2009 Cohort	2012	294
	2010 Cohort	2012	350
	2011 Cohort	2012	348
	2012 Cohort	2012	325

Another way of visualizing this data is to look at when the baseline and follow-up data was collected for the yield and the best practice samples (see table 3).

Table 3: Structure of Yield and BP data

Yield Data (monthly)

Year	C2009	C2010	C2011
2009	No (Baseline)	-	-
2010	Yes	-	-
2011	Yes	Yes	Yes (Baseline)
2012	Yes	Yes	Yes

Best Practice Data (twice per year)

C2009	C2010	C2011
No (Baseline)	-	-
Yes	Yes (Baseline)	-
Yes	Yes	-
Yes	Yes	Yes

The design of TechnoServe's M&E strategy is based on the premise that old and new cohorts should resemble each other on average; and that therefore subsequent Cohorts (that have not yet been trained) can be used as control regions for current Cohorts (that are already undergoing the training). The assumptions underlying this thinking include the fact that: (i) every year Cohorts are selected based on the same set of criteria; (ii) Cohorts are not geographically concentrated and are composed of 6 or more cooperatives – therefore, on average, we would expect these Cohorts to be relatively similar, as geographic and other differences would cancel out; and (iii) farmers self-select into the program, so it is likely that the type of farmer that self-selects into the program in 2009 would have relatively similar characteristics, on average, to farmers that self-select into the program in 2010, 2011, 2012, etc.

b. <u>Current impact estimates</u>

A simple comparison of average coffee-tree yields for farmers in year 1 or 2 of the training program with baseline yield data for farmers that were just starting the training (see table 4), suggests that TechnoServe's agronomy program has had a significant impact on yields in participating cooperatives (see table 5). Not only are yield levels in trained Cohorts higher than in non-trained Cohorts, but the more years of training a Cohort has undergone, the greater the average productivity of coffee trees (see table 4). The differences between yields in 'Cohorts under training' and 'new Cohorts' are statistically significant. Assuming that all Cohorts were quite similar on average before the start of the program (given their geographic composition and the selection criteria), a first estimate of the potential impact of the Agronomy program on coffee yields is 30-35% in year 1 of training for Cohorts 2010 and 2011 (there seems to have been little impact on Cohort 2009 in year 1), and an additional 10% in year 2 (see table 5). TechnoServe refined these estimates by eliminating outliers with excessively high reported yield levels, and reported 51% impact for Cohort 2009 after two years of training and a 52% impact for Cohort 2010 after one year of training.

Table 4: Average yields (kg/tree) in program Cohorts

Cohorts / years	2010	2011	2012
Cohort 2009	2.04 (Year 1)	3.04 (Year 2)	3.23 (Post-training)
Cohort 2010	1.99 (Baseline)	2.86 (Year 1)	3.13 (Year 2)
Cohort 2011	N/A	2.10 (Baseline)	2.85 (Year 1)
Cohort 2012	N/A	N/A	2.21 (Baseline)

Table 5: Estimated cumulative impact of training on average yields (kg/tree)

Impact	Cohort 2009	Cohort 2010	Cohort 2011
Year 1	3%	36%**	29%**
Year 2	45%**	42%**	N/A

***statistically significant at the 1% level*

Increasing yields in participating cooperatives seems to have been accompanied by a rapid increase in best practice adoption. According to best practice data collected from a random sample of participants who attended at least 50% of sessions in year 1 of training, 92% of farmers adopted at least 50% of best practices in 2010, a number that increased to 96% in 2011.

Partly as a result of the large impact observed, as well as potential biases inherent in the approach, TechnoServe decided to commission this study to re-assess the accuracy of its impact measures and the internal validity[2] of the treatment and control groups (i.e. subsequent Cohorts).

c. Potential biases

There are number of potential biases affecting both the yield and best practice estimates. These can be grouped into the following 4 categories:

1. ***Selection bias***: There is a risk of selection bias at a number of levels:

   a. There was no randomization at the Cohort or cooperative level, which means farmers were not randomly assigned to either the treatment or control group. Even though selected cooperatives and program participants are likely to be relatively similar on average (see explanation above), the lack of randomization implies that this is not necessarily the case. Moreover, TechnoServe does not currently have enough socio-economic indicators on individual farmers to prove how similar or dissimilar the Cohorts are on average. Structural differences between Cohorts could therefore explain some of the differences we observe between the treatment groups (Cohorts that have received training) and the control groups (Cohorts that are just about to start the training and for which baseline data is available).

   b. Cooperatives that make up the 2009 Cohort are different in nature from cooperatives in the 2010, 2011 and 2012 Cohorts. The first set of cooperatives TechnoServe worked with in the 2009 Cohort were existing cooperatives with a history, while cooperatives in the 2010, 2011 and 2012 were relatively new. Moreover, in terms of the yield samples, the 2009 Cohort yield sample was selected amongst farmers that had attended a certain number of sessions and adopted a certain number of best practices, whereas farmers in the yield sample in Cohorts 2010, 2011 and 2012 were randomly selected at the beginning of the program. These differences make the 2009 Cohort non-comparable to the 2010, 2011 and 2012 Cohorts.

   c. The "best practice" samples in each Cohort were constructed by randomly selecting farmers that had attended at least 50% of sessions in year 1. This means that there is no information on the impact of the project on farmers with less than 50% attendance in year. We therefore only have information on a sub-sample of the training population, which limits the applicability of the results to the entire training population.

[2] An experiment has **Internal Validity** when the setup minimizes the risk of biases (e.g. selection bias, confounding factors, enumerator bias, etc.) thereby making it possible to establish causality.

2. ***Missing observations in time.*** Missing observations in time affect both the yield and best-practice adoption analysis:

   a. In terms of yields, the main constraint we face is that there is only one observation in time for the control region, i.e. baseline data on yield levels in new cooperatives is only collected once, at the beginning of the training period. The problem with having only one baseline observation in time is that it is impossible to conduct a typical difference-in-difference analysis. We discuss ways around this problem in the ensuing sections.

   b. In terms of best practice adoption, the constraining factor is that no baseline data was collected for Cohorts 2009 through to 2012. This means that while it is possible to gain useful insights from the link between best practice adoption and other variables such as farmer attendance, it is not possible to infer impact. Attempts to identify a valid instrument or to estimate impact through matching were not successful.

3. ***Over-reporting of results and enumerator bias.*** There are two important aspects to this:

   a. One of the main issues with respect to TechnoServe's M&E system is that "Farmer Trainers" themselves have been responsible (up until Cohort 2012) for collecting the data rather than independent enumerators. Given that they also provide the training, Farmer Trainers could in theory have an incentive to over-state results to show that they are delivering on the job. This potential over-reporting by Farmer Trainers could explain why the program has recorded such impressive best practice adoption rates over the past four years. (The counter argument to this, and we will show that there is a lot of validity to this perspective as well, is that monitoring best practice adoption in the field is a technical task and cannot simply be conducted by enumerators with no agronomy training. Training enumerators to do this is expensive and not straightforward.) Moreover, farmers will have developed a trust relationship with the Farmer Trainers that they do not have with the enumerators leading to potential respondent related biases.

   b. Another source of bias is the nature of the best practice data that TechnoServe collects. We can illustrate this with a simple example. One of the indicators data collectors have to report on is whether the coffee leaves in a given plot look healthy or on the contrary yellow. Green leaves indicate good health and tree nutrition, while yellow indicates disease. The data collectors (either the Farmer Trainers or enumerators) have to make their own assessment as to whether the leaves in a given coffee farm/plot look green or yellow enough on average to justify a tick or a cross. This leads to enumerator bias.

4. ***Indicator design.*** TechnoServe collects data on 11 best practices, summarized in table 6 below. For certain indicators, the way results are reported can lead to an over-estimation of actual adoption rates. This is particularly an issue for three indicators: rejuvenation (BP6), safe use of pesticides (BP8), and erosion control (BP 10). Currently, farmers who do not need to rejuvenate their coffee trees because they are too young are counted as adopters of the rejuvenation best practice; farmers who do not use pesticides are counted as adopters of safe pesticide use; and farmers who do not need to worry too much about erosion because their land is flat are counted as adopters of erosion control measures.

Table 6: Best Practice Indicators

Indicator	Description
BP1	Record Keeping
BP2	Mulching
BP 3	Weeding
BP 4	Nutrition
BP 5	Composting
BP 6	Rejuvenation
BP 7	Pruning
BP 8	Safe Use of Pesticide
BP 9	IPM
BP 10	Erosion Control
BP 11	Shade Management

Chapter 3. Methodology & Overview of Findings

To assess the impact of the Agronomy program on coffee tree productivity and best practice adoption we focused on four complementary tools:

- Internal validity checks to identify potential biases to the identification strategy using additional data (in particular soil and climatic data, rainfall, additional socio-economic indicators, etc);
- Field spot-checks on a sample of 300 farmers to check the accuracy of reported best practice adoption data, combined with additional data collection ;
- Specific strategies to overcome identified biases and answer the question of whether the program had an impact on coffee tree yields and best-practice adoption in participating cooperatives; and
- Semi-structured interviews and focus groups with coffee farmers to test some of the initial findings in the field.

Detailed methodologies on how we went about estimating the impact of the program are developed in detail in the next two chapters: the first on the impact of the program on yield levels, the second on the impact of the program on best practice adoption. A short summary of the approach we used for each is provided below:

a. Strategy for estimating the impact of the program on coffee tree productivity

We follow a three step process to determine whether there is sufficient evidence to support the claim that the Agronomy program had an impact on average yield levels:

- ***We first study potential biases,*** in particular: (i) evidence of selection bias by comparing Cohorts under training (the treatment group) to new Cohorts (the control groups) using a new dataset, containing information not only on yields, but also cooperative level data (including altitude, number of trees, soil and climatic conditions, etc), sector level data from the coffee census, and district level socio-economic data; (ii) issues relating to omitted variables in time; (iii) potential over-reporting by farmer trainers and (iv) the effect of annual coffee cyclicality on yield levels.

- ***Second we propose a strategy to overcome some of the identified sources of potential bias***, by: (i) adjusting for distributional differences between Cohorts and reducing the risk of over-reporting by focusing on a sub-sample of farmers; (ii) estimating the time effect in order to overcome the fact that we only have one observation in time for the control groups; and (iii) controlling for observed differences at the individual, cooperative, sector and district levels. This strategy enables us to provide an estimate of the impact of the coffee program on a specific sub-sample of farmers, which is representative of more than 70% of program participants. We eliminate the effects of annual coffee cyclicality by looking at results over a two-year period, i.e. when the coffee trees are back to their original "high-yield" or "low-yield" state.

- ***Finally, we seek to establish a link between yield increases and best practice adoption.*** While this link does not prove causality, it is a strong indication that increased best practice adoption is potentially what led to yield improvements.

b. Strategy for determining whether the program had an impact on best practice adoption

Given the structure of the data and in particular the lack of a baseline, we are not able to provide a formal estimate of the impact of the program on best practice adoption. Instead, we build a compelling evidence-based case as to why the program is almost certain to have had an impact on best practice adoption. To do this we follow a three step process:

- ***We first we check for selection bias*,** by comparing the composition of the best practice sample to the population of program participants. Given that the best practice sample was selected amongst beneficiaries that attended at least 50% of sessions in year 1, major differences in the composition of the two groups would indicate that the best practice sample is not representative of the training population, but just a sub-sample thereof.

- ***Second we check the validity of reported best practice adoption rates, testing for Farmer Trainer over-reporting.*** We do this in two ways: (i) first by comparing Farmer Trainer reporting to reporting by TechnoServe enumerators; (ii) second by studying the results of our independent field spot checks. Spot checks were conducted on a randomly selected sample of 270 farmers in 9 identified "high-risk" cooperatives, where we observed some inconsistencies in Farmer Trainer reporting.

- ***After checking for biases, we build the case for the link between training and best practice adoption, by comparing attendance rates and trends to best practice adoption patterns.*** A strong link between attendance and best practice adoption patterns is not proof of causality going from training to adoption, but is a strong signal that this might indeed be the case. We look at this link over time and study it at the aggregate and disaggregated level.

c. Overview of results

Key findings on yields, best practices and issues related to M&E system are summarized below:

Impact of the of program on coffee yields – see Chapter 4

We find strong evidence of a positive impact of the training program on coffee yields. For a sub-group of farmers (that are representative of more than 70% of the training population), **we estimate that the Agronomy training was associated with a 72.6% increase in average yield levels for farmers in Cohort 2010 after two years of training**. Initial results suggest that the impact of the training program on Cohort 2011 could be even larger: after one year of training yield levels in Cohort 2011 had increased by 75.5% compared to 57.5% in Cohort 2010, even though these initial estimates are biased upwards by the annual cyclicality of coffee trees. An innovative technique based on the similarity of best practice adoption patterns between pairs of farmers also shows that yield increases are intricately related to best practice adoption. While the results are clearly indicative of a causal impact, it is important to note that these impact estimates might be slightly over-stating the actual impact of the program, as it is not possible to distinguish between the effect of the training program and the additional effect of monitoring (an unexpected finding of this study - see below).

Impact of the program on best practice adoption – see Chapter 5

We find a strong association between attendance and best practice adoption in all Cohorts, which is further evidence of the potential impact of the training program on best practice adoption. **The higher a farmer's attendance rate, the more likely he/she will adopt a best practice.** Supporting the hypothesis that the training program leads to higher adoption, we find that as the training progresses the difference in adoption rates between "trained" farmers and "untrained" farmers increases. Moreover, there appears to be a clear link between attending a specific training session on a certain best practice, and adopting the corresponding best practice. These quantitative findings are supported by anecdotal evidence from the field: for all best practices, more than 50% of farmers interviewed as part of a random spot-check claimed have acquired these best practices through TechnoServe's training program.

Verification of data collection methods and biases – see Chapter 4,5 &6

The design of TechnoServe's M&E strategy is based on the premise that old and new cohorts should resemble each other on average and that therefore subsequent program Cohorts of farmers (that have not yet been trained) can be used as control regions for current Cohorts (that are already undergoing the training). We identify six key risks and sources of potential bias in TechnoServe's M&E system:

1. ***The risk of selection bias due to the way successive Cohorts and M&E samples were constructed***. We find selection bias at three levels: (i) there are small but significant differences between Cohorts on both farmer and cooperative level indicators; (ii) the nature of cooperatives that make up Cohort 2009 is different from Cohorts 2010, 2011 and 2012, leading to Cohort 2009 being dropped from the analysis; and (iii) the best practice sample is not representative of the training population, but rather a sub-sample of farmers that attended 50% of sessions in year 1. We tackle selection bias by controlling for farmer, cooperative, sector and district-level indicators, including both socio-economic and agro-climatic controls, and by working with a sub-sample of farmers.

2. ***Omitted observations in time***. The fact that we do not have yield baseline data for the control groups makes it impossible to conduct a typical difference-in-difference impact analysis and to check the internal validity of the evaluation by comparing baseline data in the control and treatment groups. To circumvent this problem, we exploit the structure of TechnoServe's yield data to estimate the effect of time and adjust the results accordingly.

3. ***Over-reporting by Farmer Trainers, who conducted both the training and the data collection***. We do not find sufficient evidence to support the claim that there has been systematic over-reporting of best practice adoption or yield levels by Farmer Trainers. A simple comparison of adoption rates reported by Farmer Trainers and TechnoServe enumerators, suggests there are no significant differences between them. Moreover, while spot-checks on a sample of 270 randomly selected farmers suggest that Farmer Trainers are over-reporting results by 20 percentage points, a deeper analysis of the data shows that the most likely explanation to this difference is not Farmer Trainer over-reporting but instead enumerator error.

4. ***Annual cyclicality of coffee trees, which affects yield levels.*** We find clear evidence of annual cyclicality patterns affecting the productivity of coffee trees. Coffee trees alternate between a "high yield" and a "low yield" state, from year to year. The state of a given coffee plot is not synchronized across farmers in a given Cohort or a given cooperative. Given that we do not know in advance whether a Cohort is predominantly in a "high" or a "low" yield state, it is difficult to isolate the effect of the natural annual cyclicality of coffee trees on yield levels from one year to the next. The only way to eliminate the effect of the annual cyclicality of coffee trees is to measure results over a two-year period, when coffee trees are back in their original "high" or "low" yield state.

5. ***Indicator design, leading to an overestimation of actual adoption.*** TechnoServe collects data on 11 best practices. For certain indicators we show that the way results are reported can lead to an over-estimation of actual adoption rates.

6. ***The effect of monitoring on the way famers experience the training program (see Chapter 6)***. One of the most unexpected and surprising results of this study is the large effect of monitoring on the way farmers experience the training program. TechnoServe's extensive and unique monitoring and evaluation system in Rwanda enables us experimentally conduct one test that even the best randomized control trials would not allow: to test the impact of the monitoring and evaluation system itself on project beneficiaries. We find that the regular monitoring of farmers in the yield sample –

which consists of bi-monthly (twice per month) interactions between the farmer and the project staff for three years – led to a 12 to 15 percentage point increase on farmer attendance rates, a 7 percentage point increase in best practice adoption and a significant increase in fertilizer usage. Similarly, monitoring on farmers in the best practice sample - which only happens twice per year - led to a temporary 5 percentage point increase in their attendance rates. These results lead to interesting questions on how to best leverage the monitoring effect to increase program outputs. Creating mechanisms to provide the 'illusion' of monitoring should be piloted and tested as they might increase farmers' attendance and adoption rates.

Chapter 4. Yield Impact Results

What is the impact of TechnoServe's Coffee Agronomy Training program in Rwanda on coffee yields in participating cooperatives?

CHAPTER SUMMARY:

In this section, we provide an estimate of the impact of the training program on coffee yields. The identification strategy is based on the premise that old and new cohorts should resemble each other on average and that therefore subsequent program Cohorts of farmers (that have not yet been trained) can be used as control regions for current Cohorts (that are already undergoing the training). However, there are a number of constraints/biases to this approach: (i) evidence of selection bias (ii) missing observations in time and (iii) potential over-reporting by farmer trainers. Given these constraints, we propose an alternate strategy that attempts to overcome some of the biases by: (i) adjusting for distributional differences between Cohorts and reducing the risk of over-reporting by focusing on a sub-sample of farmers; (ii) estimating the time effect, which we cannot calculate using a simple in difference in difference because of missing temporal observations; and (iii) controlling for observed differences at the individual, cooperative, sector and district levels.

Based on this approach, we find strong evidence of a positive impact of the training program on coffee yields. For a sub-group of farmers (that are representative of more than 70% of the training population), **we estimate that the Agronomy training was associated with a 72.6% increase in average yield levels for farmers in Cohort 2010 after two years of training**. Initial results suggest that the impact of the training program on Cohort 2011 could be even larger: after one year of training yield levels in Cohort 2011 had increased by 75.5% compared to 57.5% in Cohort 2010, even though these initial estimates are biased upwards by the annual cyclicality of coffee trees. An innovative technique based on the similarity of best practice adoption patterns between pairs of farmers also shows that yield increases are intricately related to best practice adoption. While the results are clearly indicative of a causal impact, it is important to note that these impact estimates might be slightly over-stating the actual impact of the program, as it is not possible to distinguish between the effect of the training program and the additional effect of monitoring (see Chapter 6).

a. Baseline coffee yields and potential impact

TechnoServe is currently the only organization in Rwanda to collect detailed data on coffee yields at the farmer level across more than 25 cooperatives. While the vast majority of available yield data in Rwanda is either based on self-reporting or administrative reporting by wet-mills, TechnoServe collects yield data by providing scales, training and a calendar to a randomly selected group of farmers in each of the program's "Cohorts". Selected farmers enter their coffee production estimates into the calendar on a daily basis during the coffee season. The data is collected by the local TechnoServe "Farmer Trainer" once per month and then sent to the TechnoServe headquarters in Kigali to be quality checked and entered into a database. At the end of the coffee season, coffee production data is aggregated at the farm-level and divided by the number of coffee trees on the farm, which are counted once per year. Estimated baseline coffee yields were about 1.76kgs per tree for the average farmer in Cohort 2010, 1.89kgs per tree in Cohort 2011, and 2.04kgs per tree for Cohort 2012 (see Table 7).

While there is not a very high degree of variation in yield averages at the cooperative level - baseline cooperative-level yield averages range from 1.25 kgs/tree in Gisuma to 2.64 kgs/tree in Cafeki – variation is very high within cooperatives. This is a direct consequence of the sensitivity of coffee trees to a whole range of factors: the quality and composition of the soil, altitude, rainfall, the age of the tree and the branches, agronomic practices (in particular weeding, mulching, pruning, rejuvenating, composting, shade management), the use or not of fertilizer in optimal quantities, the incidence of pests and diseases, the socio-economic situation of the farmer and the prevailing production cycle (coffee production in Rwanda has relatively strong on-years and off-years). It also reflects the fact that coffee trees within a cooperative are not always aligned in their annual coffee cycle: some coffee trees are in a "high yield" state, while other are in a "low yield" state (note that coffee trees alternative between years of high and low productivity). Within each cooperative baseline yields ranged from about 200-400 grams per tree to between 5-7kgs per tree, with a standard deviation in most cases of more than 1kg per tree.

Table 7: Baseline Yields in 2010, 2011 and 2012 Cooperatives

Cooperative	Cohort	Baseline Yield (kg/tree)			
		Average Yield	Minimum Yield	Maximum Yield	Standard Deviation
Cafeki	2010	2.64	0.11	6.78	1.40
Gisaka	2010	1.76	0.24	5.75	1.00
Giseke	2010	2.79	0.12	5.88	1.67
Gisuma	2010	1.25	0.23	4.81	1.15
Musha	2010	2.01	0.17	7.96	1.58
Mwezi	2010	1.36	0.07	4.30	1.08
Karama	2011	2.18	0.60	4.55	0.83
Kinyaga	2011	1.82	0.23	7.46	1.56
Koakagi	2011	1.44	0.13	6.66	1.32
Matyazo	2011	1.98	0.09	5.48	0.96
Nasho	2011	2.08	0.32	6.97	1.24
Shara	2011	1.59	0.66	7.53	1.46
Vunga	2011	2.29	0.25	5.65	1.14
Bwishaza	2012	1.58	0.55	5.19	1.32
Gasange	2012	2.20	1.10	5.76	1.32
Gishyita	2012	1.93	0.53	4.82	0.87
Kigembe	2012	1.57	0.35	3.23	0.82
Mayaga	2012	1.54	0.55	5.09	1.37
Mayogi	2012	2.29	0.60	3.27	0.75
Mukindo	2012	1.87	0.67	4.65	1.18

The potential impact of TechnoServe's training program, based on the underlying agronomic principles, is very large. TechnoServe runs a number of demonstration plots in participating cooperatives to test yield potential: the demonstration plots consist of a selected number of trees (about 40) on the plots of "Focal Farmers" and are used for training purposes. Production on the "demonstration plots" is controlled and combine best agronomic practices with the right type and dosage of fertilizers and pesticides. For the 2010 Cohort on which we have complete data, the "demonstration plots" yielded 6.82kgs/tree on average in 2011, compared to a baseline of 1.75kgs/tree for the Cohort on average. This corresponds to a potential increase in yields of 228%.

In the next section we propose a strategy to estimate the potential effect of the program on coffee yields. Although we cannot fully isolate the program effect, the consistency of the results across Cohorts 2010 and 2011 strongly suggests that the program is having an impact on yield levels.

b. Proposed strategy to estimate impact of agronomy program

While the design of the agronomy program and the corresponding data collection strategy do not provide for a robust experimental set-up that enables causal inference, we adopt an **alternative** strategy to estimate the impact of the training on coffee yields. The strategy is designed to overcome some of the main elements that threaten the internal validity of this exercise, namely: (i) the fact that there was no randomization in the assignment of individuals and/or cooperatives to the "treatment groups", which results in selection bias; (ii) missing observations in time, making it impossible to do a typical difference-in-difference estimation of impact; (iii) potential over-reporting by farmers or "Farmer Trainers"; (iv) the annual cyclicality of coffee trees, which experience high-yield years and low-yield years; and (v) the effect of monitoring on the way farmers experience the program.

Despite the biases there are a number of elements in TechnoServe's M&E set-up that can be used to generate a robust estimate of the impact of the program on coffee yields in participating cooperatives. The consistency with which cooperatives have been selected, farmers have registered into the program, and "Yield" samples have been constructed - starting with Cohort 2010 - makes it possible for us to propose an alternative strategy to estimate the impact of the program on coffee yields. The proposed strategy attempts to overcome some of the biases, by: (i) adjusting for distributional differences between Cohorts and reducing the risk of over-reporting by focusing on a sub-sample of farmers; (ii) estimating the time effect, which we cannot calculate using a simple in difference in difference because of missing temporal observations; and (iii) controlling for observed differences at the individual, cooperative, sector and district level. We eliminate the effects of the annual cyclicality of coffee trees by measuring results over a two year period.

c. Analysis of biases that threaten internal validity of identification strategy

In this section we focus on four potential biases that threaten the internal validity of the exercise aimed at estimating the impact of the TechnoServe agronomy program on yield levels: selection bias, omitted observations in time, potential over-reporting by farmer trainers and biases related to the annual cyclicality of coffee trees. Another confirmed bias - that we call the "monitoring effect" - is explained in detail in Chapter 6.

(i) Selection bias

The Coffee Agronomy program was not designed as an experiment whereby farmers, cooperatives, or both were randomly assigned to treatment and control groups[3]. Farmers in selected cooperatives self-selected into the program by turning up at the first two/three sessions and by registering their names. All registered farmers qualified to participate in the training, which means that there is no control group of registered farmers that did not receive the training within a given Cohort. The treatment group is in effect the entire Cohort, which means that the only possible control group is another Cohort with similar characteristics.

The problem from an experimental point of view is that cooperatives that met the minimum criteria were not randomly assigned to one Cohort or another. Cooperatives are ranked every year based on a set of criteria:

[3] TechnoServe has altered its approach to monitoring and evaluation in 2012: (i) farmers in Cohort 2012 have been randomly assigned to treatment and control groups; (ii) BP and Yield baseline data has been collected for the 2012 Cohort, which was not the case previously; and (iii) independent enumerators have replaced farmer trainers for data collection.

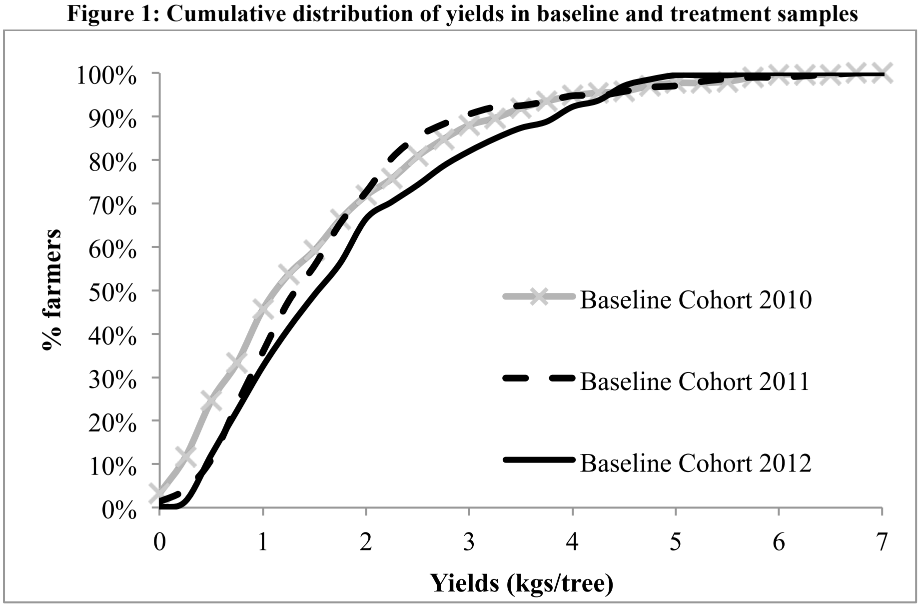
the cooperatives with the best fit are selected into the program, while the others are left out. On the one hand, this means that TechnoServe's successive Cohorts are likely to be quite similar on average, given that they are made-up of cooperatives that were selected using the same set of criteria and consist of farmers that self-selected in the program; on the other hand, it could mean that the most similar cooperatives have already been assigned to one Cohort and that the next batch will be somewhat different. The absence of randomization at the cooperative levels implies that successive Cohorts are likely to be slightly different on average, possibly resulting in selection bias.

Selection-bias: small but statistically significant differences between Cohorts

We test the extent to which this lack of randomization resulted in observable **selection bias**, by comparing the 2010, 2011 and 2012 Cohorts on yields as well as alternative characteristics at the individual, cooperative, sector and district levels. While the Cohorts appear to be quite similar on average, the differences between them are nevertheless statistically significant on multiple indicators.

TechnoServe has collected baseline yield data for the 2010 Cohort in 2010, the 2011 Cohort in 2011 and the 2012 Cohort in 2012. Even though these baselines were captured in different years and vary from 1.76kgs/tree in Cohort 2010 to 2.05kgs/tree in Cohort 2012 (a difference of about 17%) we find no statistically significant difference on average between them. This holds when controlling for individual farmer-level characteristics, cooperative level topological characteristics, sector level coffee production indicators and/or district level socio-economic indicators. These results suggest that before the intervention the 2010, 2011 and 2012 Cohorts had similar coffee yields on average and that the effect of time on coffee yields in selected cooperatives appears to be not significant.

Figure 1: Cumulative distribution of yields in baseline and treatment samples



A closer look at the distribution of yields within Cohorts however brings to light some small but nevertheless significant differences. Figure 1 compares the cumulative distribution of baseline yields in Cohorts 2010, 2011 and 2012. A simple Kolmogorov-Smirnov equality-of-distributions test enables us confirm whether these baseline yield distributions are very similar or not[4]. We find that while we cannot reject the null

[4] The Kolmogorov-Smirnov test is a nonparametric test for the equality of continuous, one-dimensional probability distributions that can be used to determine if two datasets differ significantly by quantifying the distance between the empirical distribution function of the sample and the cumulative distribution function of the reference distribution.

hypothesis that the baseline distribution of yields is identical in the 2011 and 2012 Cohorts, the distribution of yields is significantly different in Cohort 2010. The difference comes mostly at the lower and higher ends of the distribution. Assuming that the impact of the agronomy program on coffee yields is different for each farmer depending on his/her starting point, these differences could lead to a biased estimation of impact. This is something we will need to control for when estimating impact.

The next question is whether these Cohorts are similar on other variables of interest, including individual farmer characteristics, topological conditions at the cooperative level, coffee production indicators at the sector level and socio-economic indicators at the district level? While the Cohorts appear to be relatively similar, the differences between them are nevertheless statistically significant on multiple indicators (see table 8). Overall the 2010 and 2011 Cohorts are quite similar: we cannot reject the null hypothesis that both Cohorts have similar sized farmer plots on average, similar shares of female registered farmers and cooperative members, consist of cooperatives with similar levels of rain-fall and lime endowments, and cover districts with similar poverty levels. The 2012 Cohort, however, is different to both the 2010 and 2011 Cohorts on most accounts. In the 2012 there is a larger share of female registered farmers combined with fewer cooperative members, cooperatives are in slightly poorer districts, with higher population density rates and lower levels of rainfall. This suggests that yields in Cohort 2012 should face higher downward pressure than yields in the 2010 and 2011 Cohorts.

Table 8: Measuring the similarity of the cohorts

Indicator	**Cohort 2010**	**Cohort 2011**	**Cohort 2012**
Individual farmer characteristics (in sub-sample)			
Number of trees	253	219	223
Female registered farmer (% total)	26.5%	26.3%	33%
Cooperative member (% total)	29.4%	26.9%	8.3%
Topology characteristics at cooperative level			
Altitude	1611m	1685m	1645m
Rain 2009	1007mm	962mm	910mm
Rain 2010	1053mm	1031mm	978mm
Rain 2011	1102mm	1068mm	1043mm
Lime	90	85	100
Coffee production characteristics at cooperative level (2009 data)			
Average number of trees (2009 Census)	252	232	146
% productive trees (2009 Census)	57%	45%	56%
Coffee area under production (2009 Census)	163ha	107ha	113ha
Socio-economic indicators at the district level (2009 data)			
Poverty (% households, district level, 2009 EICV)	48.8%	48.7%	55.1%
Secondary education (net enrollment, 2009 EICV)	18.8%	17.8%	16.2%
Under 5 mortality (per 1000 children, 2009 EICV)	93	103	98
Density (inhabitants, per square kilometer)	383	455	445

(ii) Missing observations in time

It is not possible to conduct a standard difference-in-difference impact estimation of the effect of the Agronomy program on yields because we only have one observation in time for the "control group" (i.e. the Cohort that has yet to undergo training). TechnoServe collects baseline yield data for each Cohort in the year that training starts; the second time yield data is collected, the Cohort will already have experienced one year of training. The difference in the performance of the two groups before and after the intervention – called the "difference-in-difference" – would enable us to isolate what change would have happened anyway because of the effect of time, regardless of the treatment, and what change was induced by the treatment. We can better visualize the problem in the difference-in-difference equation below, where T and C represent the treatment and controls groups and the index represents time:

$$Impact_1 = (T_1 - T_0) - (C_1 - C_0) \quad (1)$$

The way the data is currently structured we can observe T_1, T_0 and C_1, but not C_0. In other words we have baseline and follow-up data for the "treatment group", but only follow-up data for the "control group". This leaves us with two problems: (i) we cannot isolate the effect of time, which is $(C_1 - C_0)$; and (ii) we cannot check whether the control and treatment regions, respectively C_0 and T_0, are similar before the training. Let's say for example that we were interested in the impact of Cohort 2010 and that we used Cohort 2011 as a control group. We would be able to observe baseline yields for Cohort 2010 in 2010 (T_0) and baseline yields for Cohort 2011 in 2011 (C_0), but we would not be able to tell whether differences in average yields between them were due to time or inherent differences between Cohorts.

(iii) Over-reporting by farmers and farmer trainers

A potential bias to the results reported in this section is over-reporting by farmers and/or farmer trainers. While we cannot formally exclude the possibility of over-reporting, we have reasons to believe that this is not the case for the yields data:

- Farmers do not have any obvious incentives to over-report production levels. While low best-practice adoption rates could potentially represent a failure on their part, Farmer Trainer have less responsibility for yield levels on a farm. Yields can be affected by a whole range of external factors. The lack of direct responsibility also means that Farmer Trainers have fewer incentives to over-report yields data.
- Farmers provide "Farmer Trainers" with a written monthly report on daily production levels, which contrasts with the way data is collected for best-practices. Best practice data is collected by the Farmer Trainer himself and entered into a database via SMS on the spot. Changing data on production levels requires changing what the farmer has reported in writing. While it is highly unlikely that any changes would go noticed given the amount of data involved, it still represents a somewhat higher effort and risk for the "Farmer Trainer".
- Production levels only make sense when compared to the number of coffee trees the farmer owns. Information on the number of trees a farmer owns are only collected once during the production season and yield levels are only calculated later by the TechnoServe monitoring and evaluation team. It is unlikely that Farmer Trainers would intentionally alter production data in order to achieve a certain yield level given all the steps involved.
- We find significant variations in yield levels in all cooperatives both before and after the training, varying from about 0.2kgs per tree to over 7kgs per tree. Standard deviations in yield levels increase as the project reaches completion. This combined with the consistency of reported averages in years 0, 1, and 2 of the project, strongly suggests that production numbers have not been significantly and consistently altered by either the farmers or the Farmer Trainers.

(iv) The annual cyclicality of coffee trees

Coffee trees, in particular when they do not receive the right mix of inputs and care, have a tendency to have very large yield fluctuations year in and year out. Coffee trees have up years and down years, suggesting that farmers with low baseline yield levels are likely to see their yields increase in the subsequent year, while farmers with very high baseline yield levels are likely to see the productivity of their coffee trees decrease. Which state a given coffee plot is in – a "high yield" or "low yield" state – is not completely synchronized across farmers in a given Cohort or even a given cooperative. Given that we do not know in advance whether a Cohort is predominantly in a "high" or a "low" yield state it is difficult to isolate the effect of the natural annual cyclicality of coffee trees on average yield levels from one year to the next.

Tracking the yield fluctuations of coffee trees belonging to program participants over time enables us to shed some light on the dynamics of this cycle. Data comparing a farmer's coffee tree productivity in one year and the next (regardless of the Cohort and the year) clearly shows this dynamic at play: low yield farmers see their yields increase from one year to the next, while high yield farmers see their yields decrease (see figure 2). Had it not been for the annual cyclicality of the coffee trees one would have concluded that the project was highly beneficial for low-yield farmers and highly detrimental for high-yield farmers.

Figure 2: Yield levels in one year compared to yield levels in the previous year

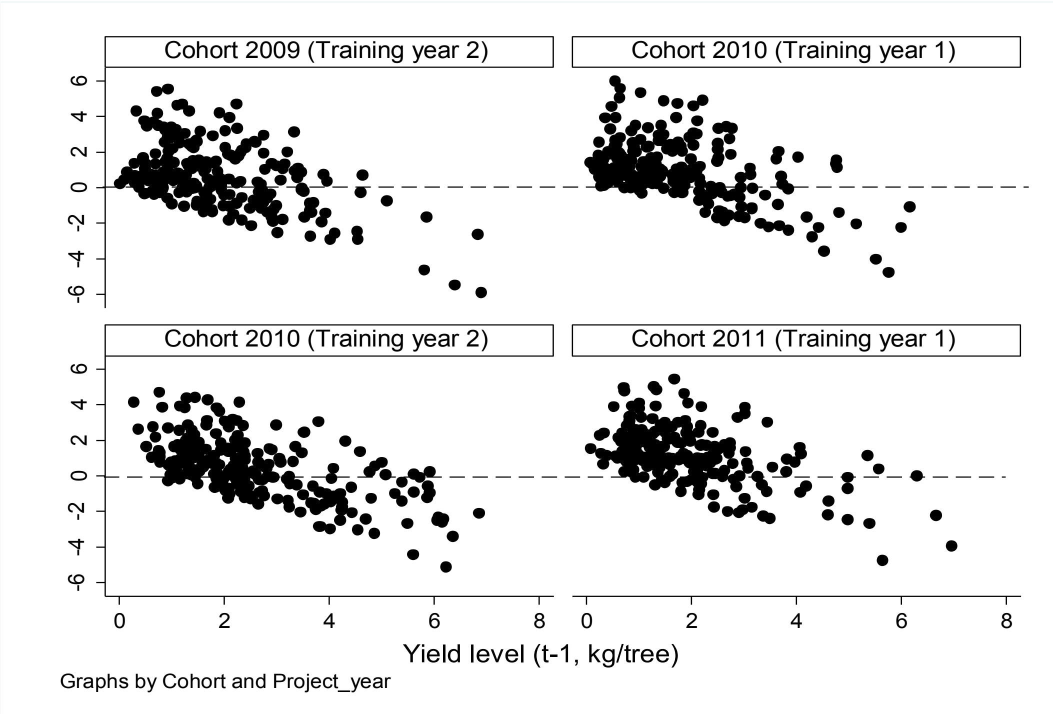
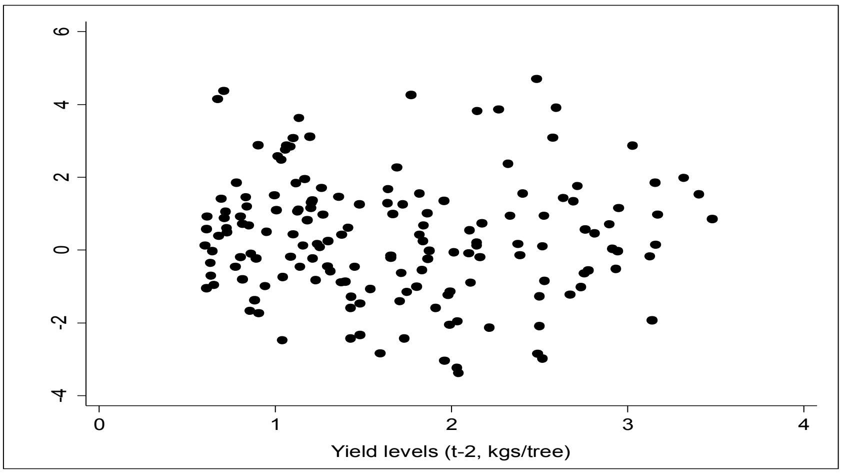


Figure 3: Yield levels in one year compared to yield levels two years earlier (Cohort 2010, t=2012)



When we look at average yield levels over a two year period though, we find that annual cyclicality cancels out. While there is a very clear negative relationship between a farmer's yield levels one year and his/her yields the next, there is no clear relationship between an individual farmer's baseline yield levels and his/her yield levels after two years. This is clearly illustrated in figure 3, which depicts baseline yield levels for farmers in Cohort 2010 and their yield levels after two years of training. These results mean that the only way to deal with biases resulting from the of annual coffee tree cyclicality is to compare baseline yield levels with yield levels after two years of training – or in other words compare the productivity of coffee trees at the same point in their annual coffee cycle. In-between observations would include not only the effects of the program, but also the effects of annual coffee cyclicality.

The implications of annual coffee cyclicality for this analysis is that it is only possible to provide an impact estimate for the 2010 Cohort for which we have 3 observations in time (a baseline in 2010, a follow-up in 2011, and a second follow-up in 2012 after the end of the training period). The 2009 Cohort has been discarded from the analysis because of selection bias, and at the time of writing we only had two observations in time for the 2011 Cohort.

To summarize, in this section we find that: (i) there is evidence that the lack of randomization has led to statistically significant differences between successive Cohorts; (ii) the fact that we only have one observation in time for the control Cohorts means that it is not possible to conduct a difference-in-difference and that alternative solutions are necessary; (iii) that even though we cannot formally exclude the risk of Farmer Trainer or farmer over-reporting, there are strong reasons to believe that systematic over-reporting with respect to yield data is unlikely; and (iv) that to deal with the inevitable biases resulting from the annual cyclicality of coffee trees, we need to compare outcomes to the baseline when the trees are at a similar point in their coffee cycle (i.e. every two years).

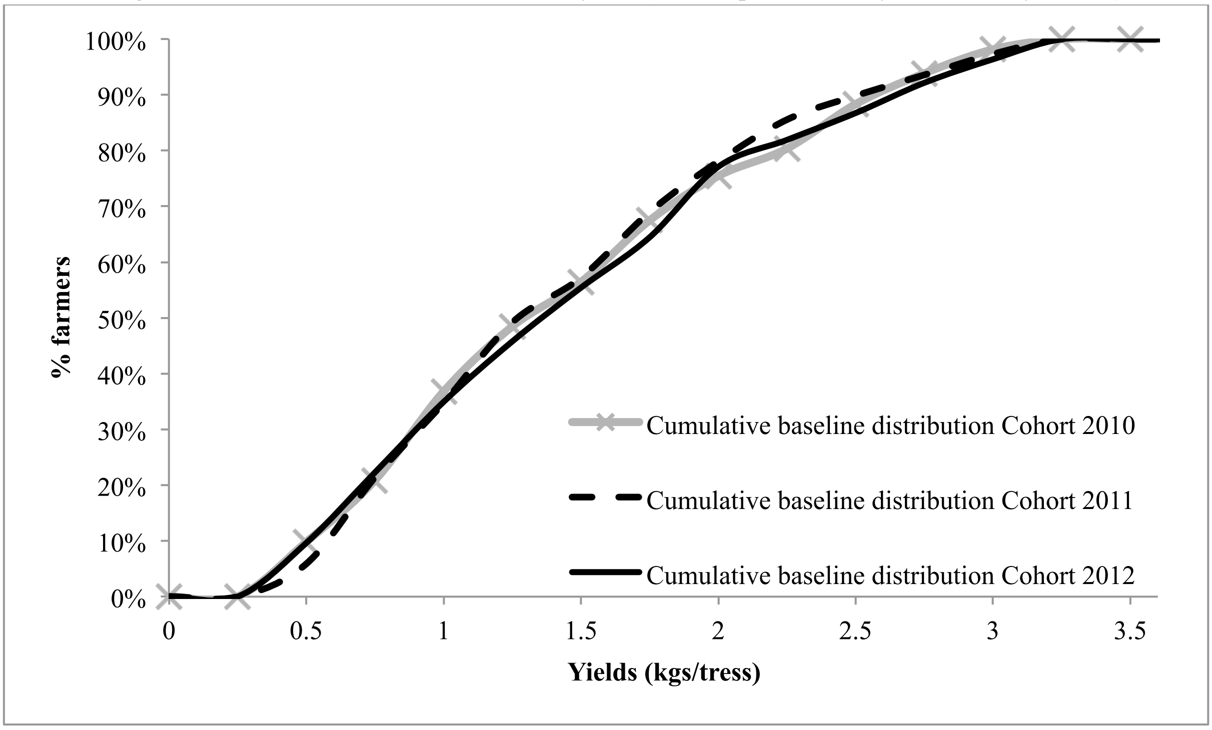
d. Estimating the impact of TechnoServe's agronomy program on coffee tree productivity

In this section we propose a strategy to overcome the identified biases and provide an estimate of the impact of TechnoServe's agronomy program on yield levels for a sub-sample of the training population. We achieve this by censoring the sample based on upper and lower bound yield thresholds, by estimating the effect of time and by using an appropriate controls mix, including individual farmer, cooperative, sector and district level indicators.

(i) Eliminating distributional effects and limiting over-reporting by working with a sub-sample of farmers

In order to adjust for the differences in yield distribution resulting from selection bias and limit the risk of over-reporting we decide to focus on a sub-sample of farmers with baseline yields between 0.6-3.5kgs/tree. For the sake of consistency and to avoid differences in averages due to attrition, we further slice the samples by eliminating farmers for which data was not collected consistently over the years (i.e. we eliminate farmers from Cohort 2010 for which we do not have 3 observations in time and eliminate farmers from Cohort 2009 for which we do not have 2 observations in time). We also eliminate from the sample all farmers who reported yields higher than 7kg/tree in any given year. As can be seen in Figure 4, once we slice the yield samples in this way, the distribution of baseline yields for Cohorts 2010, 2011 and 2012 become remarkably similar despite the fact that data was collected in subsequent years, thereby removing any distribution effect from the observed metrics. The average yields for this sub-sample of farmers is 1.691kgs/tree in Cohort 2010, 1.670kgs/tree in Cohort 2011, and 1.705kgs/tree in Cohort 2012, and these average values are not statistically significant different from each other. This sub-sample, which is representative of more than 70% of farmers, consists of the main target group for the agronomy project: i.e. farmers that do not already have extremely high yields who in all likelihood might already be implementing some of the best practices, but also farmers who are starting from a minimum base and who might be able to benefit more fully from the training provided.

Figure 4: Cumulative distribution of baseline yields (sub-sample limited to yield>0.6 and yield<3.5)



(ii) Estimating the effect of time

Table 9 lists the average yield data obtained for Cohorts 2010, 2011, and 2012 in our sub-samples between 2010 and 2012. In the ensuing sections we use the numbers provided in parenthesis as a reference to the specific cell. Baseline data for each Cohort is colored light red, while follow-up data is in light green.

Table 9: Average Yields data for 2010, 2011 and 2012 Cohorts in Sub-Sample

Cohorts / years	2010	2011	2012
Cohort 2010	1.691 (1)	2.755 (4)	3.073 (7)
Cohort 2011	N/A (2)	1.670 (5)	2.899 (8)
Cohort 2012	N/A (3)	N/A (6)	1.705 (9)

If subsequent Cohorts were good controls for each other and all the required data was available, we would calculate the impact of the Agronomy program on Cohorts 2010 using the following difference-in-difference estimations, where numbers in parenthesis correspond to the cells in table 4 (note that we only estimate impact over a two year period in order to eliminate the effect of annual coffee cyclicality which affects results from one year to the next):

$$Impact_{2010} = [(7) - (1)] - [(9) - (3)]$$

The first part of this difference-in-difference estimation – which corresponds to $[(7) - (1)] -$ refers to the observed change in the treatment group (i.e. Cohort 2010) after two years of training. The second part– corresponding to $[(9) - (3)]$ – refers to the observed change in the control group (Cohort 2012) over the same period of time. The change in the control group is basically the effect of time, or in other words what would have happened regardless of the training. The problem is that (3) is unobservable, as no data was collected for Cohort 2012 in 2010.

We attempt to overcome this problem by showing that the effect of time on coffee yields in the selected Cohorts is likely to be negligible. To test this hypothesis it is necessary to assume that selection biases in the attribution of farmers and cooperatives to Cohorts 2010, 2011 and 2011 are small enough such that after controlling for key characteristics, the main difference between baseline yields is the effect of time. Despite the inherent selection biases in the project set-up, this is a reasonable assumption to make. There is an adequate level of internal validity in the way subsequent Cohorts were created, which explains why the difference between Cohorts on selected baseline indicators is small. Factors that have contributed to enhancing internal validity include:

- The fact that at the cooperative level there are many potential candidates for TechnoServe's coffee agronomy program. To select its first batch of 6 cooperatives for example, TechnoServe ranked 88 different cooperatives based on the selection criteria described above. The greater the number of cooperative-candidates, the more likely it is that cooperatives selected in subsequent years will be similar on average. If there were only 18 cooperatives in total in Rwanda and TechnoServe had selected 6 in Cohort 2009, 6 in Cohort 2010 and the remaining 6 in Cohort 2011, it is unlikely that the 2009 batch would resemble the 2011 batch in any way; with a pool of more than 80 cooperatives to choose from however, the opposite is true.

- Cooperatives have been selected consistently using the same set of criteria, including geographic and topological characteristics, their coffee production status, and cooperative management practices. This means that cooperatives in consecutive Cohorts should be more similar to each other on average than any other combination of cooperative-candidates that were excluded from the program because they did not meet the criteria.

- In practice, the program's current Cohorts are composed of 6 to 10 cooperatives with a similar geographic distribution in the country. The fact that there are 6 to 10 cooperatives and not 1 or 2 in each Cohort, means that at the aggregate level some of the idiosyncratic cooperative-level differences are averaged out.

- Farmers within these Cohorts and cooperatives should be comparable, as they have self-selected into the program, thereby demonstrating a certain interest in the program and a willingness to learn. Moreover, the yield data captured in the yield sample is based on a representative sample of registered farmers, randomly selected amongst farmers that registered and attended the first training session.

If this assumption holds, after controlling for differences in key characteristics between Cohorts, we should obtain:

$(5) = (1) + t_1 + \in_1$ where $\in_1$ is negligible and t_1 is the effect of time in 2010-2011; and,

$(9) = (1) + t_1 + t_2 + \in_2$ where $\in_2$ is small and t_2 is the effect of time in 2011-2012.

The difference between (9) and (5) should therefore give us an estimate of t_2, while the difference between (5) and (1) should give us an estimate of t_1.

Controlling for individual, cooperative, sector and district level characteristics, we find that we cannot reject the null hypothesis that: $t_1 = t_2 = 0$, i.e. we cannot prove that t_1 and t_2 are not 0. Table 10 summarizes the results we obtain:

Table 10: Regression results estimating impact of time

	Coefficient	Interval	Controls
Estimate t_1	-0.04kgs/tree	[-0.24;0.12]	Individual characteristics, topological characteristics at cooperative level, coffee status at sector level, socio-economic indicators at district level
Estimate t_2	-0.09kgs/tree	[-0.42;0.24]	

We can further test this estimate of t_2 using the difference between (8) and (4), which corresponds to follow-up data for Cohorts 2010 and 2011, i.e. one full year into the training program. Using the same notation for the effect of time and letting P_1 be non-time related changes to yield levels in Cohort 2010 during 2010-2011 (including the program effect and the effect of annual coffee cyclicality) and P_2 non-time related changes to yield levels in Cohort 2011, we can write:

$(8) = (1) + t_1 + t_2 + P_2 + \omega_2$ where ω_2 is very small; and,

$(4) = (1) + t_1 + P_1 + \omega_1$ where ω_1 is very small.

Re-arranging these terms we obtain:

$t_2 = [(8) - (4)] - (P_2 - P_1) + \omega_3$ where ω_3 is very small.

We estimate P_2 by taking the difference between (8) and (9), and estimate P_1 by taking the difference between (4) and (5). Controlling for all the relevant indicators, we find that we cannot reject the null hypothesis that $t_2 = 0$. The regression results yield a point estimate of 0.11kgs/tree, with a 95% confidence interval ranging from a lower bound of -0.25kgs/tree to an upper bound of 0.47kgs/tree. While this point estimate is slightly higher than the previous estimate of t_2 the conclusion remains valid.

(iii) Selecting the right control variables

Equipped with an estimate of the effect of time it is now possible to provide an estimate of the how the program has affected coffee yields. Maintaining our previous assumption on the scale of the selection bias, we can estimate the impact of training on Cohort 2010 by measuring the difference between follow-up data for Cohort 2010 (the treatment group) collected in 2012 and baseline data for Cohort 2012 (the control group) and adjust for the effect of time ($t_1 + t_2$). By measuring impact over a two year period we reduce biases related to the annual cyclicality of coffee trees. Formally this would translate into the following equations:

$$Y_t = \beta_1 Cohort_{t-2} + \beta_j Control_j + \varepsilon_1 \text{ where } \varepsilon_1 \text{ is very small; followed by:}$$

$$Impact_{Cohort\ t-2} = \beta_1 - t_{[t-2;t]} + \varepsilon_2 \text{ where } \varepsilon_2 \text{ is very small}$$

where Y_t are yields in year t; $Cohort_{t-2}$ is a dummy for belonging to Cohort *t-2* (the treatment) as opposed to Cohort *t* (the baseline); $t_{[t-2;t]}$ is the effect of time on yields during the [*t-2; t*] period; $Impact_{Cohort\ t-2}$ is our estimate of impact for $Cohort_{t-2}$; $Control_j$ accounts for the various controls, and the betas are regression coefficients. See Annex B for detailed regression results.

<u>Selecting the most appropriate controls</u>

Before proceeding to estimate the impact of the training on Cohorts 2010, we need to establish a list of valid controls. To correct for selection biases we control for baseline differences between Cohorts at multiple levels of spatial aggregation: differences in basic individual farmer characteristics, topological differences at the cooperative level, sector-level differences in the intensity of coffee production, and differences in socio-economic characteristics at the district level. In the selection of controls, we are constrained by sample size, the number of available indicators and the risk of multicollinearity:

- ***At the individual level***, we can only distinguish farmers in baseline and treatment samples based on the following characteristics: number of coffee trees owned, gender, and whether they are members of a cooperative or not. We do not have any other baseline data on individual farmers such as age, income, education, distance from training site, other sources of income, altitude of coffee farm, composition of soil, etc. Given that the spatial distribution of cooperatives is relatively consistent across Cohorts we have little reason to believe that there will be major difference in average individual characteristics. We nevertheless add cooperative, sector and district level controls to account for any major socio-economic differences at the aggregate level.

- ***At the cooperative level***, TechnoServe has captured topological data, which accounts for the average altitude, rainfall, lime endowments and the timing of the coffee season in each cooperative. These are factors that stay constant over time. Cooperatives are also grouped into 6 different types of "topologies". We chose however to only include topological data on altitude and rainfall in the regression analysis, as including additional indicators leads to multicollinearity. Including dummies to account for aggregate cooperative level effects or including a dummy on which topology a certain cooperative belongs to, leads to jumps in the impact estimates because both cooperative dummies and topology dummies perfectly predict Cohorts, which is our variable of interest. Being in one cooperative or another for example perfectly determines which Cohort you belong to; likewise, having a certain type of rare topology for example perfectly predicts whether you are in Cohort 2010, 2011 or 2012. We also chose to exclude other indicators, such as lime endowments or the timing of the coffee season from the list of controls, as they are closely related to altitude and rain and hence do not add predictive value. The continuous nature of the altitude and rain measures enables us to control for important cooperative level differences without creating too much multicollinearity.

- ***At the sector level***, which almost perfectly overlaps with cooperative level data (there is only 1 cooperative per sector in Cohorts 2010, 2011 and 2012), we have baseline data collected during the 2009 Coffee Census on coffee production and the age of coffee trees in the sector. This data enables us to control for initial differences in the intensity of coffee production in the areas covered by TechnoServe's successive Cohorts before the start of the training. In particular we chose to control for the share of trees classified as "productive" in the Census (aged 3 to 30 in 2009). In line with the approach for cooperatives, we opt for a continuous variable rather than sector dummies to control for these spatial differences, given that sector dummies would perfectly predict which Cohort a farmer belongs to.

- ***District level data***, collected in EICV 3 (2009)[5], is required in order to control for socio-economic differences between Cohorts before the start of the program. This is the most disaggregated level of baseline socio-economic data available. On average there are about 2 cooperatives per district, spread out across Cohorts 2010, 2011 and 2012. Some districts are only represented in certain Cohorts, which means that it is again not possible to control for district level dummies without generating multicollinearity. We chose to use the average poverty rate at the district level to proxy for differences in socio-economic characteristics in the program's various Cohorts. Other data, such as education levels, income, or health related indicators, are closely related.

This combination of controls provides us with a mix of individual, topological, socio-economic and spatial controls, which in theory should capture a large share of the baseline differences between Cohorts within our sub-sample of farmers. Missing from the regression analysis are omitted individual characteristics (age, income, etc.) and spatial dummies, which do not work because they closely predict which Cohort a farmer belongs to. Given that farmers were randomly assigned to the yield samples and that we control for spatial differences on socio-economic indicators, there is little reason to believe that there are large differences in individual farmer characteristics across Cohorts. In order to compensate for the lack of spatial dummies, which from the cooperative through to the province levels create too much multicollinearity and lead to jumps in the estimates, we cluster standard errors at the cooperative level.

(iv) Estimating the impact of the program on coffee tree productivity for a sub-sample of farmers

For farmers with baseline yield levels ranging from 0.6kgs/tree to 3.5gks/tree, **we estimate that participating in the Agronomy program was associated with an increase in yields of 72.6% after two years of training for farmers in Cohort 2010.** That corresponds to an increase of 1.24kgs/tree for the average farmer in the Cohort 2010 sub-sample. The results, which are summarized in table 11 are statistically significant and assume that: $t_1 + t_2 = 0$. The fact that yield levels have increased significantly after two years of training signals that the increase observed in year 1 for Cohort 2010 was not due to the annual cyclicality of coffee trees. Had the program had no impact at all, we would have expected yield levels to drop to their low baseline equilibrium. It is important to note however that these estimates to not distinguish between the program effect and the additional effect of monitoring, which we show in chapter 6 increased attendance rates for farmers in the yield sample by about 12 percentage points. We are not able to quantify the additional impact of monitoring on yields, but this impact is limited to the incremental impact of higher attendance rates in the "yields sample" vs. the training population. See Annex B for detailed regression results.

[5] Integrated Household Living Conditions Survey in Rwanda (EICV 2009)

Table 11: Impact Estimates 2010

		Impact Estimate 2010	
Controls	**Corresponding Variables**	**Point estimate**	**95% confidence interval**
None	Treatment dummy	77.36%	[50.0% ; 104.4%]
Individual	Number of coffee trees, gender dummy, cooperative member dummy	78.7%	[52.6% ; 104.3%]
Cooperative	Altitude, rainfall 2010	73.09%	[48.1% ; 97.5%]
Sector level	Share of productive trees in sector in 2009	72.9%	[47.7% ; 97.6%]
District level	Baseline poverty rate (% population)	72.6%	[37.5% ; 107.2%]

Another way to analyze this data using the same sub-sample is to look at the individual level differences in yield levels over time within the 2010 Cohort, controlling for a number of individual level characteristics (collected for the 2010 Cohort only), rather than comparing yield levels in one Cohort to another. We can estimate by how much the coffee tree productivity of individual farmers increased for Cohort 2010 by comparing their baseline yield levels collected in 2010 to their post-training yields collected in 2012, after two years of training.[6] Controlling for baseline yield levels, individual farmer characteristics (including the number of coffee trees the farmer owns, age, distance from the training site, and level of education[7]) and which cooperatives farmers belong to, we find that on average yield levels for farmers in Cohort 2010 increased by about 1.26kg/tree or 74.5% (see Technical Annex for detailed results). These figures are very similar to our impact estimate of 1.24kgs/tree or 72.6%.

At the cooperative level we find a high degree of variation in impact estimates. Table 12 lists before-training and after-training yields at the cooperative level. In all cases the change after two years of training is positive, with impact estimates ranging from 23.7% in the case of the Gisuma cooperative to 173.8% in the case of the Giseke cooperative. As can be seen in the table, the increase in yields in the first year of the program seems to have been greater than the increase in the second. There are two possible explanations for this: (i) the first is that the effect of the program could be greater in the first year of training; (ii) the second and more likely explanation, is that annual coffee cyclicality drives yield levels up in the first year and down in the second. The reason this would be the case is that we're working with a sub-sample of farmers that have a lower average yield level than the entire population (we're focusing on farmers with baseline yields between 0.5-3.5 kgs/tree). The sub-sample is under-representative farmers in a "high-yield" state with higher downward pressure on the productivity of their coffee trees, and on the contrary over-representative of farmers in their "low yield" state with upwards pressure on the productivity of their trees. Year 1 impact estimates therefore over-estimate the impact of the program as they also include the effect of annual cyclicality.

[6] Besides Cohort 2009, which we exclude from the analysis because of selection bias, Cohort 2010 is the only Cohort for which we have three observations in time).

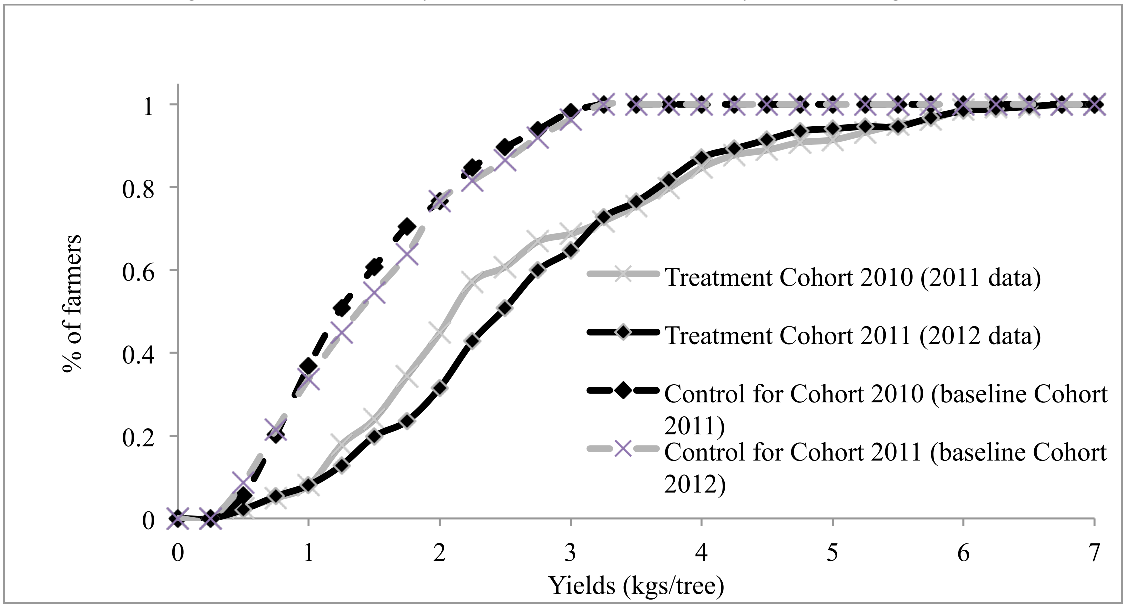
[7] This data was collected by TechnoServe for 138 yield farmers in Cohort 2010 only.

Table 12: Before- and after- training yields estimates in sub-sample

Cooperative	Cohort	Before training	After Y1 of training	After Y2 of training	Change after Y1	Cumulative change after Y2
Cafeki	2010	1.65	2.55	2.74	55.0%	62.5%
Gisaka	2010	1.49	2.61	2.72	74.8%	79.0%
Giseke	2010	1.29	3.36	3.78	161.3%	173.8%
Gisuma	2010	2.12	2.90	2.53	36.8%	23.7%
Musha	2010	1.59	3.20	3.32	100.8%	104.6%
Mwezi	2010	1.73	2.27	2.44	31.8%	39.3%

Initial results on a sub-sample of farmers with baseline yield levels between 0.5-3.5 kgs/tree suggest that the impact of the Agronomy program on Cohort 2011 could be even larger than the impact of the program on Cohort 2010. After one year of training, average yield levels in the 2011 Cohort increased by 75.5% compared to just 57.5% in the 2010 Cohort (controlling for the same set of individual, cooperative, sector and district level characteristics). Note that the 57.5% estimate for Cohort 2010 after one year of training is similar to the initial results obtained by TechnoServe, i.e. 52%.[8] As can be seen in Figure 5, the distribution of both baseline and follow-up yield levels were very similar for Cohorts 2010 and 2011, even though the increase in yield levels was slightly larger for Cohort 2011.

Figure 5: Distribution of yield data before and after one year of training



e. Establishing a link between yields and best practice adoption

Can we link the increase in coffee tree productivity in Cohorts 2010 and 2011 to greater best practice adoption? While we cannot prove that an increase in best practice adoption caused the increase in yield levels (given that we do not have baseline data for the best practice sample), we show that yield increases are intricately tied to better farming practices. We do this using an innovative methodology, which is based on a measure of how similar the farming practices of a pair of farmers are.

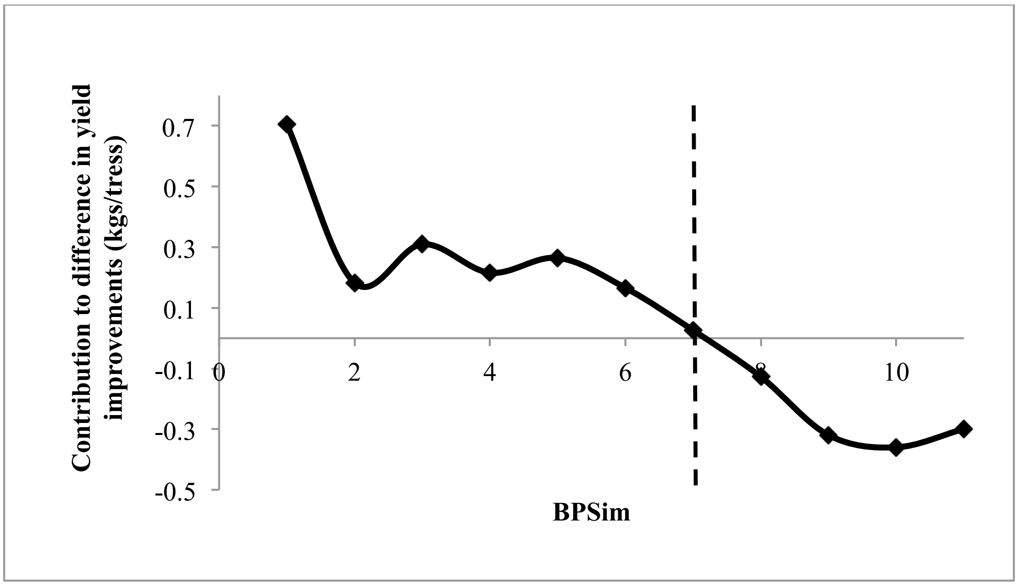
[8] TechnoServe, "*Yield Impact Memo*", 28 November 2011.

The logic of the proposed approach is relatively straightforward. In trying to identify a link between best practice adoption and improvements in yield levels, typical regression analysis will fail to deliver because what drives yield growth is not one best practice in particular, but a combination of best practices (here we have 11 best practices). With 2048 different possible combinations of these 11 best practices, one would need a very large sample to test how different combinations of best practices have led to an increase or not in yield levels. Furthermore, a simple regression including dummies for each of the best practices would not capture the interaction between best practices but instead how each individual best practice contributes to coffee tree productivity.

To overcome this issue we posit the following: if better farming practices really impact yield levels, then all else equal, farmers that utilize similar farming techniques should also experience similar increases or decreases in the productivity of their coffee trees. To test this hypothesis we propose a basic measure of how similar the best practice adoption patterns of a pair of farmers are, by simply counting the number of best practices that both farmers are implementing. E.g. if the only best practices that two farmers are both implementing are mulching and weeding, then their best practice similarity score (which we call BPSim) will be 2. We calculate BPSim for all possible pairs of farmers in Cohort 2010 (for which we have final best practice data), which is based on available data and amounts to 25,520 observations.

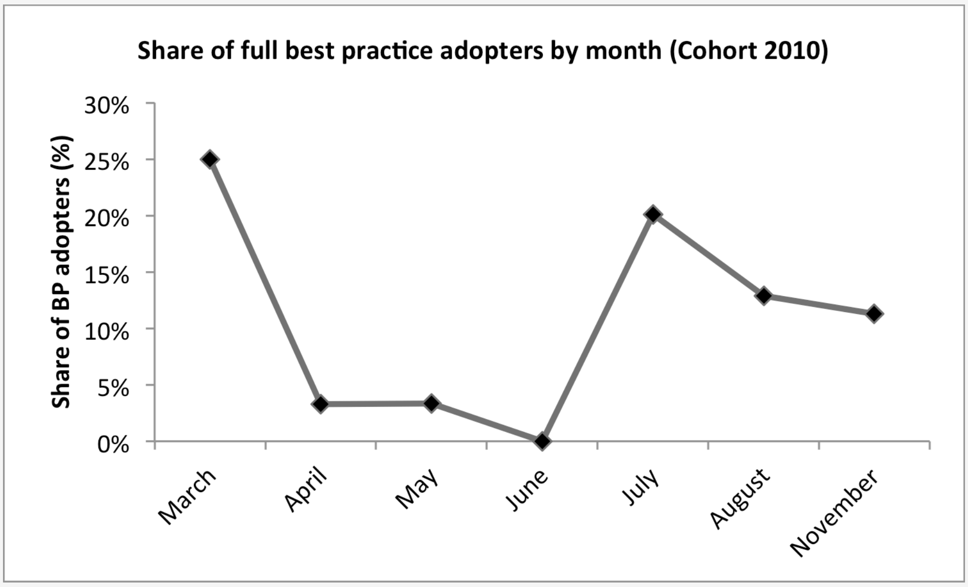
Using this basic metric we find that farmers that had adopted more similar coffee farming practices also experienced more similar improvements in yield levels after one year of training. In particular, we find that: (i) a best practice similarity score of 7 or less is associated with greater differences between yield improvements of pairs of farmers after one year of training; whereas (ii) a best practice similarity score of more than 7 is associated with more similar improvements in yield levels after one year of training (see figure 6). **This link indicates that yield improvements and best practice adoption are related**. These findings are statistically significant, controlling for the respective cooperatives of the two farmers, their gender, whether they are cooperative members or not, their baseline yields and the number of trees they own. This finding reinforces the notion that the training program has had an impact on yields through best practice adoption.

Figure 6: Relation between BPSim and the differences in yield improvements between pairs of farmers (Cohort 2010)

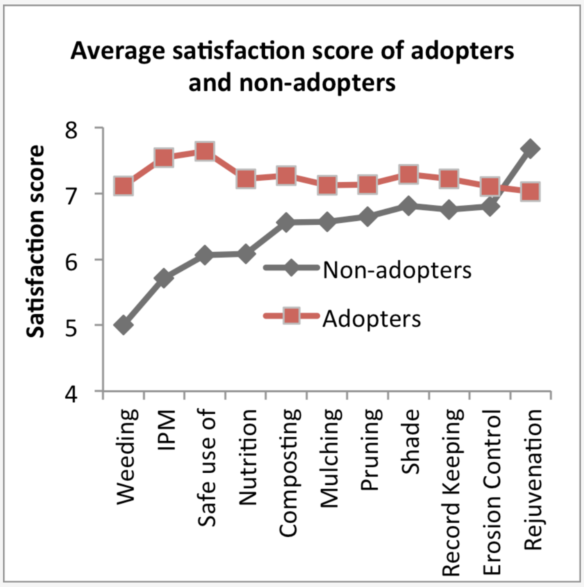


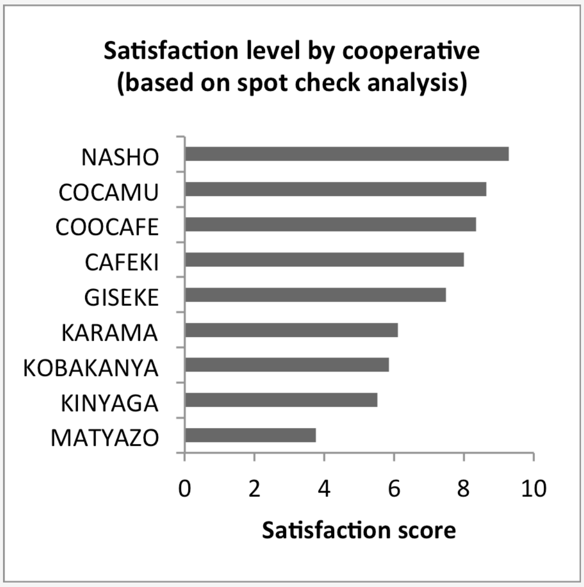
Did You Know?

- **Did you know that it is highly unlikely to find farmers that have adopted all best practices in moths other than March and July?** In March and July, more than 20% of Cohort 2010 farmers had adopted all best practices, compared to around 10% or less in any other month.



- **Did you know that adopters of best practices are much more satisfied with the program than non-adopters?** One potential explanation (out of several) is that farmers that got the most out of the program and implemented what they learned are also the most satisfied with the training. The average satisfaction score for the program based on our spot check analysis was about 7 out of 10. Out of the 9 cooperatives interviewed, farmers in Nasho, Cocamu and Coocafe were the most satisfied of the sample, while farmers in Matyazo, Kinyaga and Kobanya were comparatively less satisfied.





Chapter 5. "Best Practice" Adoption Results

Is TechnoServe's coffee agronomy training program achieving its objective of teaching and convincing participant farmers to adopt best practices? And if so, by how much have adoption rates increased?

CHAPTER SUMMARY:

In this section, we show that there is a strong evidence linking attendance to best practice adoption. Given the structure of the data and in particular the lack of a baseline, we are not able to provide a formal estimate of the impact of the program on best practice adoption. Instead, we build a compelling evidence-based case as to why the program is almost certain to have had an impact on best practice adoption. To do this we follow a three step process: (i) first, we check for selection bias; (ii) second, we check the validity of reported best practice adoption rates - testing for Farmer Trainer over-reporting, in particular through field spot checks on a randomly selected group of farmers in "high risk" cooperatives; and (iii) lastly, build the case for the link between training and best practice adoption, by comparing attendance rates and trends to best practice adoption patterns.

We find a strong association between attendance and best practice adoption in all Cohorts, which is further evidence of the potential impact of the training program on best practice adoption. **The higher a farmer's attendance rate, the more likely he/she will adopt a best practice.** Supporting the hypothesis that the training program leads to higher adoption, we find that as the training progresses the difference in adoption rates between "trained" farmers and "untrained" farmers increases. Moreover, there appears to be a clear link between attending a specific training session on a certain best practice, and adopting the corresponding best practice. These quantitative findings are supported by anecdotal evidence from the field: for all best practices, more than 50% of farmers interviewed as part of a random spot-check claimed have acquired these best practices through TechnoServe's training program.

Establishing a clear link between best practice adoption and the training program will add one more layer to the argument that there is a direct impact between the training program and the observed increase in yields. In this chapter we study best practice adoption rates, the structure of available data and potential biases, and establish a link between adoption and training.

a. Data

TechnoServe collects "best practice" data from a selected sample of registered farmers in each cohort to track progress on adoption rates on 11 measurable best practices, as well as on the usage of various nutritional products, insecticides and the incidence of major pests and diseases. The program reports annually on the share of farmers that have adopted 50% or more of these best practices (i.e. at least 6 best practices out of 11). These include: mulching, weeding, pruning, rejuvenation, erosion control, shade management, composting, nutrition, integrated pest management, safe use of pesticides and record keeping. Most of these best practices are directly observable in the coffee farms, i.e. the data collector can see whether there is mulch under the canopy or not, whether there are weeds in the field, can determine whether there are any major nutritional deficiencies based on the color of the leaves, or check whether the trees have been pruned, how old the stems are, whether there is enough shade in the field and look for signs of erosion control. The only questions that rely on self-reporting are questions related to the use of pesticides, knowledge of integrated pest management, and the application of fertilizers.

Understanding the characteristics of the best practice data is critical to developing a robust analysis of the link between adoption and training. Firstly, no single best practice is significantly more important than all the others in increasing the productivity of the coffee trees. For example, a farmer can weed and prune satisfactorily but if he/she does not provide the right nutritional mix, the trees can succumb to nutritional deficiencies, pests and diseases. Secondly, best practices are not only inter-linked but are also intricately inter-related: for example, mulching is one way of managing erosion; composting and mulching are part of nutrition; shade management and mulching lead to fewer weeds; optimal nutritional practices reduce the incidence of pests and are a central component of integrated pest management, etc. This has a number of implications for this analysis: (i) conceptually it is important to look at these best practices as a package of inter-related indicators where the law of the least common denominator applies – the number of farmers that adopt more than 50% of best practices could be 90%, but one best practice where the adoption rate is only 25% can reduce the effect of all the other best practices combined: (ii) technically, given that these best practices are not independent of each other, it becomes somewhat more complicated to isolate the effect of the program on one best practice or another.

b. Reported best practice adoption rates

Table 13 summarizes reported best practice adoption rates for farmers in Cohorts 2010 and 2011 that attended at least 50% of training sessions in year 1 (we refer to the latter as "trained farmers"), compared to a random sample of registered farmers in Cohort 2012 in the first few months of the program (Cohort 2012 is the only cohort for which we have baseline best practice adoption data on registered farmers). To ensure the results are consistent we use "round 1" data for Cohort 2010, collected in the first half of 2011, and "round 1" data for Cohort 2011, collected in the first half of 2012. We also exclude from the sample farmers that belong to the "yield sample" as they were selected using a different set of criteria.

Table 13: Reported Best Practice Adoption Data

Indicator	Description	Cohort 2010 (2011 data)	Cohort 2011 (2012 data)	Baseline Cohort 2012 (2012 data)
BP 1	Record Keeping	88%	44%	0%
BP 2	Mulching	84%	94%	72%
BP 3	Weeding	94%	96%	62%
BP 4	Nutrition	92%	85%	21%
BP 5	Composting	68%	64%	38%
BP 6	Rejuvenation	84%	95%	91%
BP 7	Pruning	82%	92%	25%
BP 8	Safe Use of Pesticide	94%	63%	43%
BP 9	Integrated Pest Management (IPM)	73%	61%	9%
BP 10	Erosion Control	93%	99%	83%
BP 11	Shade Management	46%	39%	25%
S 50	Share with more than 50% of BPs	97%	92%	34%
S 75	Share with more than 75% of BPs	65%	47%	0%
S 100	Share with all BPs	20%	19%	0%
	Observations	781	794	435

Reported results reveal remarkably high adoption levels after 1 year of training in Cohorts 2010 and 2011, suggesting the program largely achieved its training objectives. The average best practice adoption rate for trained farmers was about 82% in Cohort 2010 and 76% in Cohort 2011, compared to only 43% in the 2012 Cohort at the very start of the training program. While more than 90% of trained farmers had adopted more than half of the best practices after one year of training in both Cohorts 2010 and 2011 (this is the main indicator the agronomy program has to report on), that figure was only 34% in Cohort 2012. The largest impact seems to have been on important best practices such as nutrition and pruning along with integrated pest management and record keeping. These are followed closely by weeding, composting and the safe use of pesticides. While the potential impact of the program seems to be large, we nevertheless observe a rapid decline in adoption rates after a certain threshold: in Cohort 2011 for example, while 92% of farmers had adopted at least 6 best practices, 47% had adopted at least 9, and only 19% had adopted the full 11. We use this variation to establish a clear link between the training and the high best practice adoption rates later in this chapter.

Anecdotal evidence from the field reinforces this hypothetical link between training and the observed high adoption rates for trained farmers. As part of the spot-check analysis, we conducted surveys and focus groups in certain cooperatives in the 2009, 2010 and 2011 Cohorts and asked farmers where they had learned the best practices from, i.e. either TechnoServe or other sources, including family, friends or other government and donor funded programs. On all best practices, more than 50% of the 270 farmers interviewed claimed to have learned these from TechnoServe's agronomy program. Here are some representative statements that farmers made during focus groups in the selected cooperatives (see Annex A to see the full transcripts of each of the focus groups):

- **Farmer from Cocamu cooperative (2009 Cohort):** *"We used a lot of traditional techniques that we had learnt from our parents and friends but it's only since TechnoServe came we got professional training".*
- **Farmer from Coocafe cooperative (2009 Cohort):** "*Honestly, most of the best practices I learnt from TechnoServe: composting, IPM, shade management … there are so many"*
- **Farmer from Cafeki cooperative (2010 Cohort):** *"We knew how to do some of the best practices, TechnoServe showed us how to do it better and make sure it worked"*
- **Farmer from Kinyaga cooperative (2011 Cohort):** "*Long time ago, the agronomy government official used to come and just tell us to do this and that but TechnoServe came all the way to the field and showed us how to do it. Now we know how to use the NPK, we know how and when to pick our berries from the coffee tree and now we get more yields from the same tree".*

c. Trends in best practice adoption

In order to understand some of the dynamics of best practice adoption, we highlight a number of relevant trends that affect adoption rates: (i) the impact of prior knowledge, and (ii) the impact of time.

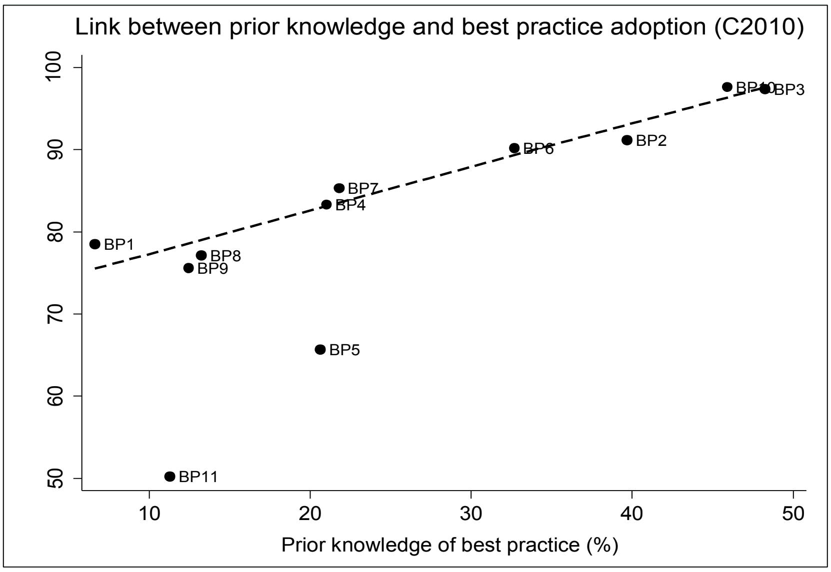
<u>(i) Impact of prior knowledge on Best Practice adoption rates</u>

Which best practices did farmers know about before the program started and how does that affect adoption rates after the training?

Based on baseline data for Cohort 2012, we find that registered coffee farmers were already - to a certain extent - mulching, weeding, rejuvenating coffee trees and managing erosion. As one farmer in the Kinyaga cooperative (Cohort 2011) put it: *"there are several methods that we learnt from our fathers and neighbours, it is just that they didn't pay attention to it and simply planted the coffee trees and left them to grow, but some of the methods included mulching and fighting soil erosion"*. Another farmer in Cafeki (Cohort 2010) claimed that "*before we got any training, we knew how to mulch, fertilize, weed, rejuvenate and harvest"*. We confirmed these baseline figures by asking farmers during the spot-check surveys where they had learned each of these techniques. The main best practices coffee farmers claim to have known about before the program started were: weeding (48%), erosion control (46%), mulching (40%) and rejuvenation (33%).

When we compare prior knowledge to actual post-training best practice adoption in Cohort 2010 (round 2 data, collected at the end of the training), we find a remarkable linear relationship. i.e. the greater the prior knowledge of a certain best practice, the higher the observed adoption rates after two years of training (see figure 7). Excluding best practices 5 (composting) and 11 (shade management) that are outliers, this result suggests that any potential program impact was almost perfectly proportional to prior knowledge about a best practice.

Figure 7: Link between prior knowledge and best practice adoption (C2010)



This measure of prior knowledge enables us to divide best practices into three groups, with composting (BP5) and shade management (BP11) as outliers[9] (see table 14). It is interesting to note that the group of best practices for which prior knowledge and post training adoption was the lowest, are record keeping (BP1), safe use of pesticides (BP8) and integrated pest management (BP9) which are the only best practices that are not directly observable in the field. Record keeping relies on being able to read and write, safe use of pesticides of having the right protective equipment and integrated pest management of remembering the different aspects of integrated pest management during the interview.

Table 14: Best Practice groups of prior knowledge

	Best Practice Group	**Corresponding best practices**
1	High prior knowledge, high post-training comparative adoption rate	Weeding (BP2), Mulching (BP3), Erosion Control (BP10)
2	Medium prior knowledge, medium post training comparative adoption rate	Nutrition (BP4), Pruning (BP7) and Rejuvenation (BP6)
3	Low prior knowledge, low post-training comparative adoption rate	Record Keeping (BP1), Safe use of Pesticides (BP8), Integrated Pest Management (BP9)

[9] One potential explanation as to why shade management is an outlier is because it is highly dependent on the presence of banana trees (or similar foliage) in the production area; a farmer could have limited knowledge of shade management but if the coffee farming area also has substantial banana trees to provide shade, then this could be counted as adoption. The case of composting is similar as it relies on post-harvest residues from other crops. If a farmer doesn't grow any other crops, then acquiring the required materials for composting can be problematic.

(ii) Impact of time on Best Practice adoption rates:

What is the impact of time on best practice adoption rates? How does seasonality affect adoption rates? Do adoption rates increase or decrease over time?

Based on reported adoption trends for trained farmers in Cohort 2010 (for which we have 3 data points in time) we distinguish 3 types of best practices:

1. Best practices that display seasonal trends (see figure 8),
2. Best practices for which adoption has been increasing over time (see figure 9); and
3. Best practices for which adoption rates have been declining (see figure 10).

The 4 best practices that showed seasonal but steady adoption rates over time were mulching, weeding, nutrition and rejuvenation. These trends correspond to the different phases of the coffee season. Figure 6 depicts adoption rates of these best practices for trained farmers in Cohort 2010 during the following periods: April to June 2011 (Round 1), July to August 2011 (Round 2) and April to June 2012 (Round 1). We observe an increase in rejuvenation, weeding and mulching during the months of July and August – i.e. Round 2 - while nutrition trends higher during the harvesting season, which lasts from March through to June. Fertilizer (i.e. nutrition) is applied twice per year, from March through to June and then in October and November, hence the slightly lower adoption rates observed in July and August. Rejuvenation and pruning occur in July and August, which explains the higher rejuvenation rates in round 2. July and August are also the driest months of the year. Fewer weeds grow during this time leading to a higher number of clean fields and subsequently higher weeding adoption rates. Finally mulching is applied during the harvesting season, from March through to June. The thickest layers of mulch can therefore be observed right after the harvesting season, in July and August.

Best practices for which adoption rates have increased steadily include pruning, shade management, erosion control and integrated pest management. Figure 7 illustrates the gradual increase in these 4 best practices over time. While erosion control adoption rates were high to start with, the program seems to have had a significant impact on pruning, shade management and safe use of pesticides over time. Based on the trends, it seems to be the case that these best practices need between 1-2 years to be fully appreciated and implemented by the farmers.

Best practices for which adoption rates have decreased over time include record keeping, composting, and integrated pest management (see Figure 10). Based on anecdotal evidence from the spot-check analysis and survey, farmers indicated that they were less likely to maintain record keeping after the training as many of them were illiterate, did not find value in the best practice or were wary of keeping records for tax purposes. For IPM, farmers indicated that it was difficult to remember all the various strategies to manage pests. Even though farmers were more aware of compositing initially, composting depends on the availability of the required materials, which are not always readily available especially when a farmer does not grow other crops.

Figure 8: Best practices with seasonal behaviour (Cohort 2010)

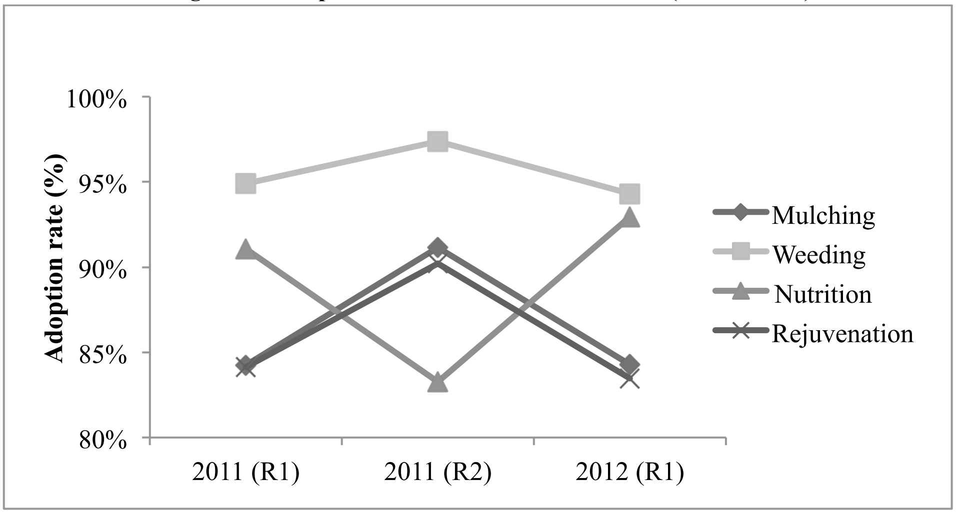


Figure 9: Best practices with increasing trends (Cohort 2010)

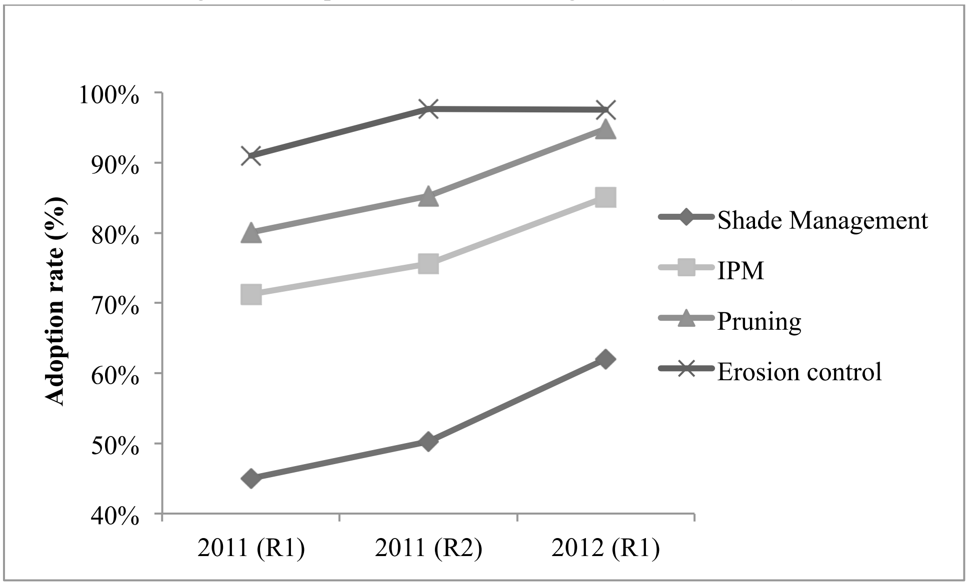
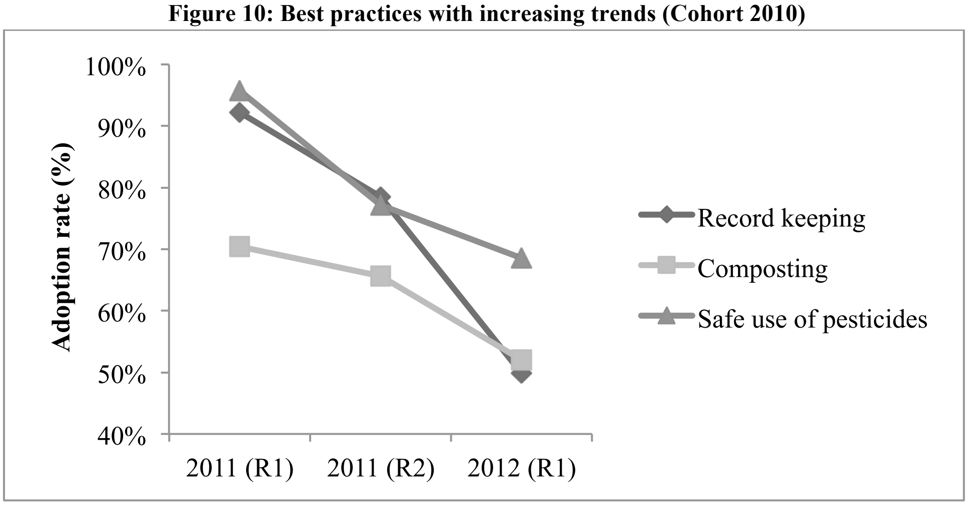


Figure 10: Best practices with increasing trends (Cohort 2010)



d. Monitoring and evaluation strategy and potential biases

Data

TechnoServe has collected best practice data for Cohorts 2009, 2010, 2011 and 2012. We focus our analysis on "best practice" data for Cohorts 2010 and 2011 only as the sample sizes in the 2009 Cohort are smaller, not consistent over time and because only baseline data is available for Cohort 2012. To understand the structure of the "best-practice" data it is useful to keep the following points in mind:

- "Best-practice" data for Cohorts 2010 and 2011 data can be divided into two distinct samples: (i) farmers that belong to the "yield sample" and for whom best practice data was also collected; and (ii) farmers that belong to the "best practice sample". The difference between the two samples is that farmers in the "yield sample" were randomly assigned to the yields sample at the beginning of the program, whereas farmers in the best-practice sample (except for the case of Cohort 2012) were randomly selected amongst farmers that had at least attended 50% of sessions after one year of training. Randomization was done at the cooperative level.
- Given that best-practice samples were only formed one year after the start of the program, there is no "best practice" baseline data available for registered farmers in Cohorts 2010 and 2011. TechnoServe does collect baseline information on agronomic practices in the selected cooperatives, but the farmers in the sample are not necessarily program participants and the metrics used are slightly different, hence not comparable.
- Best-practice data is collected twice per year, at the beginning and end of the coffee season in order to capture seasonal differences. We therefore have 3 observations in time for Cohort 2010, for which data was collected in early 2011 (round 1, 2011), late 2011 (round 2, 2011) and early 2012 (round 1, 2012). We have one observation only for Cohort 2011 (2012, round 1).
- Best-practice data for Cohort 2010 was collected by "Farmer Trainers" in round 1 and round 2 of 2011, but by independent data collectors in round 1 2012. "Farmer Trainers" also collected round 1 data for Cohort 2011 in early 2012.

Table 15 below summarizes the structure of best-practice data. The column "Yield" represents the number of farmers in the yield sample for which best practice data was collected; column "BP" the number of farmers in the best practice sample, while "(FT)" means the data was collected by a farmer trainer and "(DC)" means the data was collected by an independent data collector.

Table 15: Structure of Best Practice Data

Year	Cohort 2009		Cohort 2010		Cohort 2011	
	Yield	BP	Yield	BP	Yield	BP
2010	241 (FT)	N/A	N/A	N/A	N/A	N/A
2011 Round 1	126 (FT)	158 (FT)	209 (FT)	755 (FT)	N/A	N/A
2011 Round 2	148 (FT)	268 (FT)	197 (FT)	832 (FT)	N/A	N/A
2012 Round 1	278 (DC)	9 (DC)	300 (DC)	69 (DC)	286 (DC)	794 (DC)

Potential Biases

We focus this section on the two main issues that potentially affect how representative the best-practice samples are:

- **First, the "best-practice" sample is a sub-sample of the training population,** as it only includes farmers that attended more than 50% of the sessions after 1 year of training. To what extent the best-practice samples lead to an overestimation of the actual effect of the program on the entire population of registered farmers will depend on how different these "high-attendance" farmers are from farmers that attended less than 50% of sessions.
- **Second, the risk of "over-reporting" of best-practice adoption by the farmer trainers**. Higher adoption rates reflect well on farmer trainers' performance and hence provide them with an incentive to over-report or to be more lenient in their assessment of whether a farmer has adopted a best practice or not.

(i) How different are "high-attendance" farmers from "low-attendance" farmers?

TechnoServe collects attendance data for all registered farmers participating in the program. This is matched to a number of personal and program related characteristics, including the gender of the farmer, whether they are members of a cooperative or not, whether they are focal farmers or just participants, and how large the size of the training group they belong to is. We compare all registered farmers with high and low attendance rates in year 1 using these variables and find consistent results across Cohorts 2010 and 2011. As can be seen in table 16, a sampling frame consisting of high attendance participants only should, in theory, result in a sample with a larger share of female registered farmers, cooperative members, focal farmers, combined with smaller training classes/groups.

Table 16: Measuring the similarity of "high-attendance" vs. "low-attendance" farmers

	Cohort 2010			Cohort 2011		
Characteristics	**Low attendance**	**High attendance**	**Full attendance**	**Low attendance**	**High attendance**	**Full attendance**
Female	20.7%	25.7%	28.3%	29.7%	31.8%	29.1%
Cooperative member	15.9%	23.2%	34.3%	7.2%	19.8%	34.2%
Focal farmer	0.5%	3.4%	8.6%	0.6%	3.2%	8.4%
Average group size	26.5	25.6	24.6	37.4	33.7	32.7

To confirm these insights we compare the population and the best practice sample using the selected characteristics of interest (see table 17). We find that while the biases go in the same direction as predicted (in all cases but one – i.e. share of female farmers in Cohort 2011), the differences between the best-practice sample and the population are in fact quite small, albeit statistically significant. This suggests that there is a high degree of overlap between the population and the sub-group consisting high attendance farmers. To check this we look at average attendance rates and find that 77.8% of farmers in Cohort 2010 had attended at least 50% of training sessions in year 1 and that 77.3% of farmers had done the same in Cohort 2011. This not only confirms that a sampling frame consisting of high-attendance farmers is representative of more than ¾ of participants, but that the best-practice samples in Cohorts 2010 and 2011 are very similar.

Table 17: Population vs. Best Practice Sample Characteristics

	Cohort 2010		Cohort 2011	
Characteristics	**Population**	**Best practice sample**	**Population**	**Best practice sample**
Female	25.1%	27.9%	31.5%	28.4%
Cooperative member	22.6%	23.5%	18%	22.0%
Focal farmer	3%	4.9%	2.8%	4.6%
Average group size	25.7	24.8	34.2	33.7

To what extent do these small differences lead to an overestimation of best practice adoption rates for the entire population? To estimate the size of the potential over-estimation, we consider three different scenarios: (i) ***scenario 1:*** the highly unlikely scenario that all farmers that have attended less than 50% of classes have an adoption rate of 0 on all best practices – this enables us to put bounds on the potential over-estimation; (ii) ***scenario 2:*** a slightly more likely scenario in which farmers with less than 50% adoption would have the same adoption rate as the baseline (we use Cohort 2012 as a reference point); and (iii) ***scenario 3:*** the most likely scenario of the three in which farmers that attended less than 50% of sessions achieve ¼ of the increase registered by the high-attendance farmers compared to the baseline. Using Cohort 2010 as an example, we find that these adjusted population values - based on conservative adoptions assumptions - show that the highest risk for overestimation is on record keeping, nutrition and integrated pest management, followed closely by pruning and safe use of pesticides (see table 18). We estimate the potential overestimation on these best-practices based on scenario 3 to be between 10-15 percentage points. The potential for over-estimation on the other best practices is minor and is contained to +/- 5 percentage points.

Table 18: Potential Over-estimation Scenarios

Best Practice	Current Adoption Estimate	Scenario 1	Scenario 2	Scenario 3	Potential overestimation (scenario 3)
Record keeping (BP1)	88.2%	67.9%	68.0%	73.0%	15.2%
Nutrition (BP4)	92.2%	71.0%	75.9%	80.0%	12.2%
Integrated Pest Management (BP9)	72.9%	56.1%	58.2%	61.9%	11.0%
Pruning (BP7)	81.7%	62.9%	68.7%	71.9%	9.8%
Safe use of Pesticides (BP8)	94.5%	72.8%	82.6%	85.6%	8.9%
Weeding (BP3)	94.5%	72.8%	87.1%	88.9%	5.6%
Composting (BP5)	67.6%	52.1%	60.7%	62.4%	5.2%
Shade Management (BP11)	45.8%	35.3%	41.1%	42.3%	3.6%
Mulching (BP2)	84.3%	64.9%	81.5%	82.2%	2.1%
Erosion Control (BP10)	92.8%	71.5%	90.5%	91.1%	1.8%
Rejuvenation (BP6)	84.4%	65.0%	86.0%	85.6%	-1.2%

(ii) Farmer Trainer over-reporting

To test the hypothesis of "Farmer Trainer" over-reporting, we conduct two tests: (i) we compare TechnoServe data collected by Farmer Trainers to data collected by an independent team of enumerators recruited by TechnoServe for Cohorts 2009 and 2010; and (ii) we compare data collected during our spot-check field survey to adoption data reported by TechnoServe. The first test compares data collected by Farmer Trainers to data collected by enumerators, while the second test compares data collected by TechnoServe in general (regardless of whether data was collected by a Farmer Trainer or not) to data collected by our independent team of enumerators.

Test 1 – Comparing data collected by Farmer Trainers and Enumerators

For individuals in the best-practice sample of Cohort 2010, best practice data was collected by Farmer Trainers in 2011 (both in rounds 1 and 2 of data collection) and by enumerators in 2012 (round 1 only). We compare adoption rates reported by Farmer Trainers to those reported by the team of enumerators, keeping in mind that differences may be due to the effect of time or the last year of training for Cohort 2010, which effectively ended in August 2011. To ensure comparability, we reduce the sample to include only individuals for whom data was collected in all three periods of interest: i.e. round 1, 2011; round 2, 2011; and round 1, 2012. We are left with a sample size of 114 farmers.

Table 19: Difference between reported FT data and Enumerator data

Best practice	Round 1, 2011 (1)	Round 2, 2011 (2)	Round 1, 2012 (3)	(2)-(1)	(3) – (1)
Safe use of Pesticides (BP8)	96%	74%	56%	-23%	-40%
Composting (BP5)	80%	63%	40%	-17%	-39%
Record Keeping (BP1)	96%	82%	70%	-14%	-25%
Mulching (BP2)	91%	95%	88%	4%	-4%
Nutrition (BP4)	95%	90%	97%	-4%	3%
Weeding (BP3)	96%	99%	100%	3%	4%
Pruning	91%	82%	96%	-9%	5%
Rejuvenation (BP6)	84%	95%	90%	11%	6%
IPM (BP9)	60%	69%	70%	10%	11%
Erosion Control (BP10)	87%	98%	99%	11%	12%
Shade Management (BP11)	54%	63%	71%	10%	18%

As can be seen in table 19, there are only 3 best practices for which reported adoption rates in round 1 2012 were significantly lower than adoption rates in round 1 2011 and that hence are at risk of over-reporting. These include: safe use of pesticides (BP8), which fell from an estimated 96% of adoption in 2011 to 56% in 2012; composting (BP5), which fell from 80% of adoption to 39% in 2012; and record keeping, which was 96% in 2010 compared to just 70% in 2012.

While we cannot formally exclude over-reporting, evidence strongly suggests that the observed reduction in adoption rates for BP8, BP5 and BP1 was due to the effect of time. The decline in adoption rates for all 3 BPs between (i) round 1 2011 and round 2 2011 and (ii) round 2 2011 and round 1 2012 is consistent and gradual, e.g. BP5 adoption rates fell by 17 percentage points between round 1 2011 and round 2 2011, and then by an additional 22 percentage points between round 2 2011 and round 1 2012; (iii) while BP1 adoption rates fell by 14 percentage points between round 1 2011 and round 2 2011, and then by an additional 11 percentage points between round 2 2011 and round 1 2012. We therefore attribute these difference to time and not over-reporting by Farmer Trainers.

Based on a simple comparison of adoption rates reported by Farmer Trainers and enumerators we do not find any evidence of over-estimation of best-practice adoption rates by Farmer Trainers.

Test 2 – Spot-check analysis

To test the accuracy of the data reported by TechnoServe we conducted a spot-check analysis on 9 cooperatives in Cohorts 2009, 2010 and 2011, where potential discrepancies were identified in terms of best practice reporting. A team of enumerators was sent into the field to independently collect best practice adoption data on a sample of 262 farmers, randomly selected amongst farmers in the "best-practice" sample of these "high risk" cooperatives. Data collection took place between August 2 and August 20th, 2012. The table below provides a summary of the selected cooperatives, the rationale for selecting them and the composition of the sample.

Table 20: Spot Check Analysis Sample

Cooperative	Cohort	Rationale for selecting Cohort for Spot-check analysis	Proposed sample size
Nasho	2011	Outlier, best performer across cohorts	27
Kinyaga	2011	High performer, reporting discrepancies	27
Matzayo	2011	Unusual adoption patterns	27
Karama	2011	Outlier, worst performer across cohorts	27
Cafeki	2010	Worst performer C2010	27
Giseke	2010	Low performer	27
Coocafe	2009	Worst Performer C2009	27
Cocamu	2009	Best Performer C2010	27
Kobakanya	2009	Outlier on pruning, composting	27
		Total	**243**

Data from the spot-check analysis was merged with the latest available data collected by TechnoServe for corresponding farmers, resulting in a sample of 202 farmers[10]. At a first glance, this data appears to indicate that on almost all indicators - be it data collected by Farmer Trainers or data collected by its team of enumerators - TechnoServe has been over-estimating best practice adoption by about 20 percentage points (see Table 21). **The average adoption rate for selected farmers was 81.5% based on data collected by TechnoServe, compared to just 59.6% according to the spot checks**. The discrepancy between reported TechnoServe data and the spot-checks is particularly large on record keeping, IPM and safe use of pesticides, for which the drop is at least 40%. Reported data and spot checks agree on mulching, rejuvenation, erosion control and shade management.

Table 21: Reported vs. Spot-Check Data

Indicator	Description	Reported data (latest available	Spot-checks
BP1	Record Keeping	71%	29%
BP2	Mulching	92%	88%
BP 3	Weeding	99%	81%
BP 4	Nutrition	88%	61%
BP 5	Composting	74%	52%
BP 6	Rejuvenation	91%	92%
BP 7	Pruning	90%	61%
BP 8	Safe Use of Pesticide	65%	20%
BP 9	IPM	75%	21%
BP 10	Erosion Control	95%	92%
BP 11	Shade Management	58%	60%
	Average adoption rate	**81.5%**	**59.6%**

Such a large estimate for potential over-reporting would be a surprising result however, because: (i) we find perfect matches between spot-checks and reported data in more than 65% of cases, indicating that in the

[10] Out of the 268 farmers interviewed, we were able to match 202 farmers; 70% matches were with best practice data collected in 2012 and the remaining 30% with data collected in 2011. The reason we were not able to match all 268 farmers is because of missing observations in the best practice dataset, in particular for farmers in Cohort 2011.

majority of cases there was no over-reporting by TechnoServe data collectors or errors by our survey team; and (ii) we find equal levels of potential over-reporting regardless of whether the reported data was collected by a Farmer Trainer or a TechnoServe enumerator: in 34.2% of the cases where a TechnoServe enumerator reported "adoption" our surveyors found "non-adoption"; similarly in 34.8% of cases where a TechnoServe farmer trainer reported adoption, our surveyors found non adoption. This suggests that Farmer Trainers and enumerators working for TechnoServe have equal incentives to over-report adoption rates, which does not make intuitive sense. TechnoServe switched to using enumerators for its data collection effort in response to criticism that Farmer Trainers had a clear incentive to exaggerate best practice adoption. Extra care was therefore taken to ensure that this team of enumerators operated independently from the Farmer Trainers. There is also no self-evident reason why TechnoServe enumerators would have an incentive to exaggerate adoption rates, which leaves us with the question of whether there is an alternative explanation for the gap in adoption rates?

There are several possible explanations as to why we might observe discrepancies between the spot check analysis and TechnoServe reporting:

- the effect of time, which as in the cases of record keeping, composting and safe use of pesticides, will have a negative impact on adoption rates;
- the impact of the training program, which in all likelihood led to an increase in adoption rates for Cohort 2011, which was still receiving training up until August 2012;
- over-estimation of adoption rates by Farmer Trainers or TechnoServe enumerators; and/or
- enumerator errors during the spot checks, confusing adoption for non-adoption and vice-versa.

Although reported TechnoServe data was collected several months up to a year and a half before the spot checks were conducted (given that for some farmers best practice data was not collected in 2012), time fails to account for the observed differences in adoption. Even when controlling for time and eliminating the best practices most affected by time (i.e. record keeping, safe use of pesticides, and composting), we find that the discrepancies between the reported data and the spot checks persist. The program effect cannot be behind the observed discrepancies, because it would have resulted in an increase in adoption rates rather than a decrease. So we are left with two potential explanations: either over-reporting or enumerator errors.

We explore the enumerator error hypothesis and find that this is the most plausible explanation as to why we observe a discrepancy between the spot-checks and the data collected by TechnoServe. There are three reasons why enumerator error is a strong possibility: (i) while our enumerators were trained on two occasions by a TechnoServe "Business Trainer" on how to collect best practice data in the field, these training sessions lasted half a day and might not have been enough for the enumerators to fully grasp all the nuances involved in distinguishing adoption from non-adoption, in particular for certain best practices; (ii) best practices 1, 8 and 9, where we observe the largest drops in adoption, are also the only best practices that are not directly observable in the field – the enumerator asks the farmer whether he/she has done A or B, and then the enumerator has to determine based on the best of his knowledge whether the farmer is an adopter or not; and (iii) adoption and non-adoption are not always a clear-cut case: if a farmer has pruned some of his trees but not all, is he an adopter or not? If a farmer keeps a record book that was last updated it in July, while the interview was conducted in August, is he an adopter or not? If a handful of trees have some yellow leaves and others don't, is the farmer an adopter of nutritional best-practices or not? The way the TechnoServe check-list is set-up leaves room for interpretation and subjectivity, which could go one way or another especially when the enumerators have not had sufficient training. Could it be the case that our enumerators were right 65.8% of time (which corresponds to the exact number of matches between the spot checks and the reported data) and wrong 34.2% of the time?

To test the enumerator error hypothesis we compare the likelihood of finding "non-adoption" in the spot checks when TechnoServe reports "adoption" (we call this $error_1$) to the likelihood of finding "adoption" when TechnoServe reports "non-adoption" ($error_2$). Everything else equal and assuming over-reporting by TechnoServe Farmer Trainers and enumerators is negligible, we would expect the two likelihoods to be equal: i.e. enumerators would be as likely to make a mistake on adoption of a certain best practice as they would be to make a mistake on non-adoption. Given that there are many more "adopters" than "non-adopters" in the sample (based on TechnoServe data, average adoption was 81.5%), the average adoption rate would decrease. Table 22 summarizes this scenario ($error_1 = error_2 = 34.2\%$) in which the actual adoption rate is 81.5% but the observed adoption rate due to error in the spot-checks is only 59.6%.

Table 22: Observed Non-Adoption vs. Adoption rates (1)

	Reported	**Observed Non-Adoption (spot-checks)**	**Observed Adoption (spot-checks)**
Adoption	81.5%	27.9% [=81.5% x error]	53.6% [=81.5% x (1-error)]
Non-adoption	18.5%	12.2% [=18.5% x (1-error)]	6.3% [=18.5% x error]
Totals (spot-checks)		**40.3%**	**59.6%**

We find an almost identical pattern when looking at the actual data (see table 23). If we calculate the error terms going from the spot-check analysis back to the reported data, we would find that $error_1 = 34.4\%$; $and\ error_2 = 33.3\%$. The match between the two error terms suggests that the enumerator error hypothesis is a possible explanation as to why we observe such a large difference in means between reported data and the spot-checks. Controlling for which cooperative a farmer belongs to, when the TechnoServe data was collected, and the best practice under consideration using a logit model, we estimate that $error_2$ increases to 39.5% and $error_1$ reduces to 32.6%. This means that spot-check enumerators would be more likely to report "adoption" where TechnoServe reports "non-adoption", rather than report "non-adoption" where TechnoServe reports "adoption". Assuming that there is no over-reporting by TechnoServe data collectors and that our enumerators have one chance in three of making a mistake, this difference between the two error terms would imply that best practice adoption was in fact higher at the time of the spot-checks, the opposite of what the average adoption rates suggest.

Table 23: Observed Non-Adoption vs. Adoption rates (2)

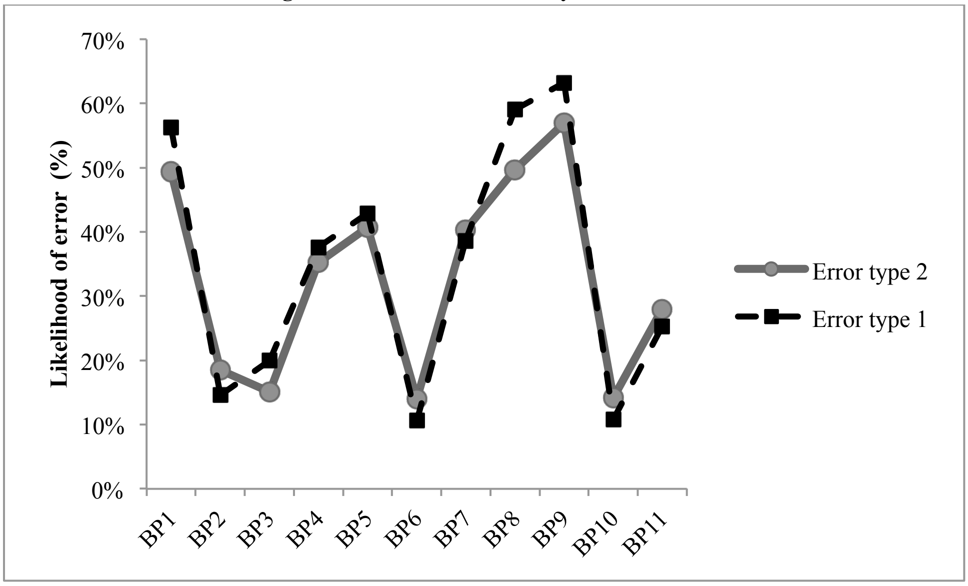
	Reported	**Observed Non-Adoption (spot-checks)**	**Observed Adoption (spot checks)**
Adoption	81.5%	28.04%	53.47%
Non-adoption	18.5%	12.33%	6.17%
Totals (spot-checks)		**40.37%**	**59.63%**

If the enumerator error hypothesis is correct, then we would expect to find equal likelihoods of mistaking "adoption" for "non-adoption" or the opposite within each best practice. If however there was over-reporting, then we would expect these two error terms to diverge depending on the best-practice. Given that there are some best practices – e.g. shade management – with many more "non-adopters" than in other best practices, over-reporting would likely be much higher in the latter. It would therefore be more likely for our surveyors to find errors of type 1 for those best practices (i.e. finding "non-adoption" when TechnoServe found "adoption") than errors of type 2 (i.e. finding "adoption" when TechnoServe finds "non-adoption"). As can be seen in figure 11, after controlling for the cooperative, the timing of TechnoServe's data collection for a specific farmer, and the best practice at hand, we find that:

- the average error rate for some best practices is much higher than in others – this could either mean that some best practices are more difficult to measure than others, that time or the training affects some best practices more, or that there is more over-reporting by TechnoServe on some best practices; and,

- within each best practice - except for BP1 (record keeping), BP8 (safe use of pesticides) and BP9 (integrated pest management) – errors of type 1 and errors of type 2 remain relatively similar, suggesting that our surveyors continue to make equal mistakes one way or the other.

Figure 11: Predicted errors by Best Practice



For BP1, BP8 and BP9, we find that the chance of our surveyors finding "non-adoption" when TechnoServe reported "adoption" is significantly higher than the alternative (i.e. $error_1 > error_2$). The effect of time on BP1 (record keeping) and BP8 (safe use of pesticides) – both tend to decrease over time – probably explains why we find that many former "adopters" have become "non-adopters" during the spot checks. We do not have any explanation however for why adoption rates decline so much in BP9, other than the fact that it is probably the most technical of the best-practices and the one where the chance of our surveyors making a mistake is consequently higher. Note that BP1, BP8 and BP9 are also the only best practices that are not directly observable in the field and the ones that depend more on the judgment and knowledge of the interviewer, which is one explanation of why error rates are the highest for those best practices.

Our final argument in defending the enumerator error hypothesis consists in looking at this data differently and assuming that the spot-check analysis is accurate, that we have no reason to doubt the results collected by the spot-check enumerators, and that any major divergence is due to errors or over-reporting by TechnoServe data collectors. This time we don't calculate the error terms by using TechnoServe data as the reference, but rather by using the spot check data as the reference: $error_1$ therefore becomes the likelihood of TechnoServe data collectors reporting "non-adoption" when in fact the farmer is adopting; and $error_2$ becomes the likelihood of TechnoServe data collectors reporting "adoption" when in fact the farmer is not adopting. In this case we would find: $error_1 = 10\%$; and $error_2 = 69\%$. This would mean that 69% of the time, TechnoServe data collectors misrepresent non-adoption and say that the farmer is adopting when in fact he isn't. Controlling for which cooperative a farmer belongs to, when the TechnoServe data was collected, and the best practice under consideration, this error term reduces to 45.6% (regardless of whether the data was collected by a Farmer Trainer or a TechnoServe enumerator) while $error_1$ increases to 26%. These error terms would suggest that: (i) TechnoServe data collectors misrepresent non-adoption almost half of the time; (ii) that Farmer Trainers and enumerators have equal incentives to exaggerate adoption; and that (iii) where there is adoption, adoption is decreasing. While we cannot formally exclude this possibility, this seems highly unlikely given that spot-check farmers reported high-levels of satisfaction with the program, that the vast majority of them claimed to have learned most best practices from TechnoServe, and that we find equal incentives for misrepresenting "non-adoption" for both Farmer Trainers and TechnoServe enumerators.

Based on the evidence presented, we conclude that the most likely scenario behind the 20 percentage point difference in average adoption between reported TechnoServe data and the spot checks is enumerator error during the spot checks rather than TechnoServe over-reporting. This finding underlines the importance of training for any future data collection efforts by TechnoServe and the need to make the check-list as specific as possible, eliminating to the extent possible the space for interpretation and human error.

e. Does training lead to higher adoption?

To determine whether TechnoServe's agronomy training program has impacted agronomic practices, we analyze the link between training and adoption using attendance as a proxy for the intensity of the training farmers have received.

Link between training and best-practice adoption

If TechnoServe's coffee agronomy training program is impacting best practice adoption, then we should be able to discern differences in best practice adoption rates between farmers that attended many sessions and farmers that did not. While this would be a strong signal that the TechnoServe training program is succeeding in changing agronomic practices in coffee producing regions, it is not in itself sufficient to establish causality. Farmers that attend many sessions for example could also be the most serious and dedicated farmers, making it impossible to distinguish between the change induced by the training program and the change induced by the individual effectiveness of the farmer.

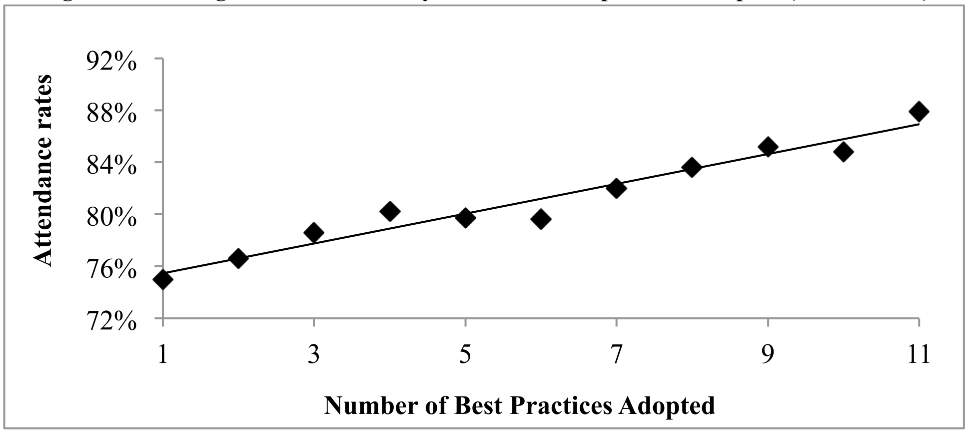
To establish a link between training and adoption we focus our efforts on famers from Cohort 2010 for which best practice adoption data was collected between July and November 2011 (i.e. round 2, 2011) at the very end of the agronomy training program for Cohort 2010. We find similar patterns for data collected in March to May 2011 for Cohort 2010 (i.e. round 1, 2011) and March to June 2012 for Cohort 2011 (i.e. round 1, 2012), but these observations suffer from too little variation in attendance rates and best practice adoption. Given that the best practice samples were selected among farmers with at least 50% attendance in year 1, variation in attendance rates immediately after the first year of training is quite limited: more than three-quarters of farmers in these best practice samples had attended more than 90% of sessions in year 1.

(i) Link between attendance and aggregate best practice adoption levels

A first glance at average attendance rates[11] based on the number of best practices adopted by farmer in Cohort 2010 (round 2), suggests that attendance and best practice adoption are highly correlated (see figure 12): the higher the average attendance rate, the higher on average the number of best practices adopted. While the incremental differences in attendance are small – there is only about a 10 percentage point difference in average attendance between low adopters and high adopters – attendance is related to an increase in the likelihood of a farmer adopting a best practice. Moreover, this relationship seems to be linear, justifying the use of linear models to compute the likelihood of adoption. This makes sense as each class relates to a different best practice: the more classes a farmer attends, the more best practices he/she learns about.

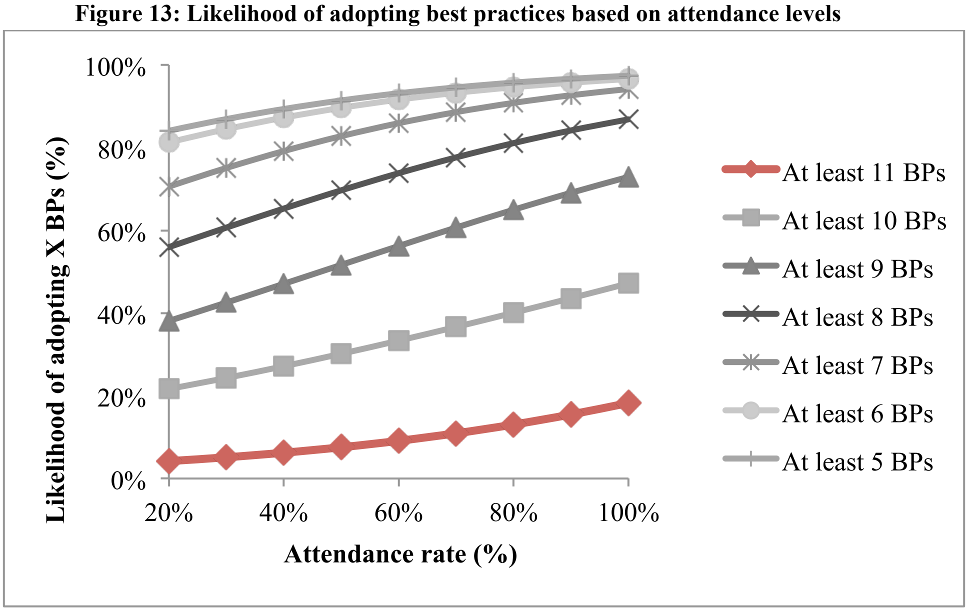
[11] We count attendance up to July 2011, before the round 2 data collection effort started. Attendance is therefore calculated based on a maximum of 16 training sessions in total.

Figure 12: Average attendance rate by number of best practices adopted (Cohort 2010)



We further explore this link between adoption and attendance by examining the likelihood of achieving a certain minimum level of best practice adoption at various attendance levels. We first study the likelihood of adopting 5 best practices at various levels of attendance, then the likelihood of adopting 6 best practices, and so forth. We calculate the likelihood using a logit model, controlling for individual farmer characteristics (including gender, whether the farmer is a cooperative member or not,), project characteristics (group size, whether a farmer participates in the program as a focal farmer, whether a farmer is also in the yield sample) and which cooperative farmers belong to. We cluster standard errors at the cooperative level. See Annex B for detailed regression results.

Figure 13: Likelihood of adopting best practices based on attendance levels



As can be seen in figure 13, regardless of the best practice threshold we select, attendance is associated with an increase in the likelihood of adoption. We estimate that the likelihood that a farmer adopts all 11 best practices if he/she has attended 20% of sessions is 4%, compared to about 18% for farmers that have attended all sessions; in the same way the likelihood that a farmer that has attended 20% of sessions adopts at least 10 best practices is 17%, compared to 47% if that same farmer has attended all sessions. In all 7 cases under consideration, the association between attendance and best practice adoption is positive and highly statistically significant suggesting that the training provided is indeed leading to higher best practice adoption rates in the field. We find similar results for the adoption of fertilizers such as lime and Zinc/Borium in particular. The increase in the reported use of composting and NPK is positive but not statistically significant.

Another way to visualize these results is to look at the evolution of the attendance rate of adopters and non-adopters over time. We do this by taking the weighted attendance rate of adopters and non-adopters per session, where we consider farmers to be "adopters" on best practices they have adopted and a "non-adopters" on the best practices they have not. In other words, using this definition, a farmer who has adopted 7 out 11 best practices is counted as an adopter 7 times and a non-adopter 4 times. We find that while it is difficult to tell apart the attendance rates of adopters and non-adopters in the first 5 months of the program (i.e. from session 1-5), the difference in attendance between them increases over time (see figure 14). This suggests that the eventual adopters and non-adopters showed the same interest in the training sessions in the first few months of the program; the growing divergence in attendance levels between them thereafter is consistent with the hypothesis that the training program is indeed leading to higher adoption rates and that the more training sessions a farmer attends the higher the likelihood that he/she becomes an adopter.

Figure 14: Attendance rate of adopters vs. non-adopters by farmers session (Cohort 2010)

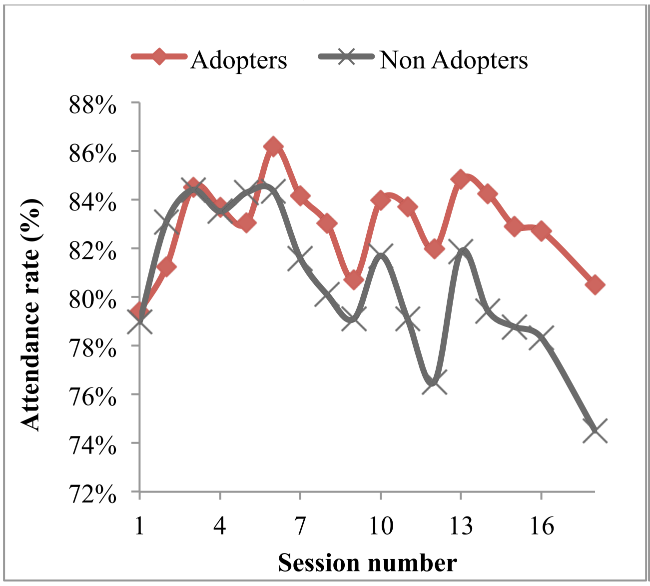
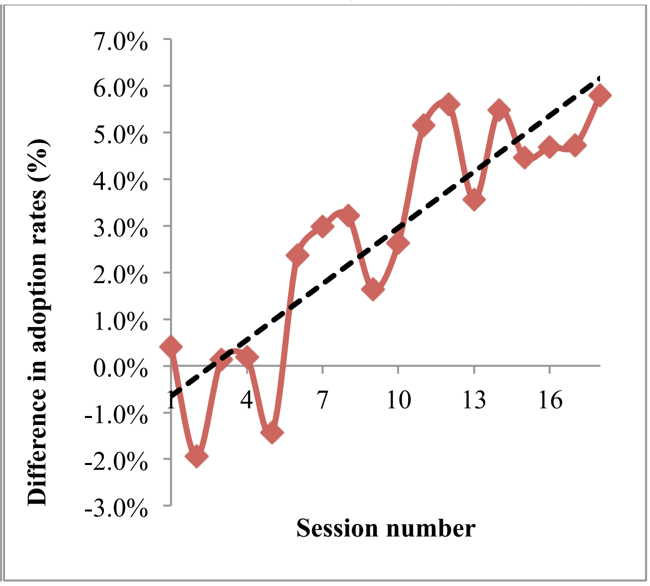


Figure 15: Difference in adoption rates between trained and untrained farmers by session (Cohort 2010)



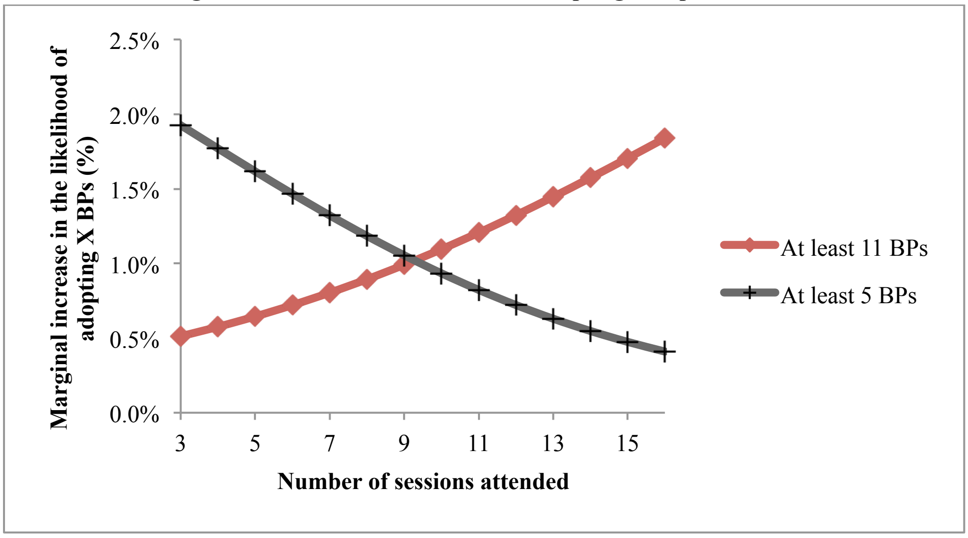
We confirm these findings by looking at the average difference in adoption rates over time of "trained" and "untrained" farmers by session – we call a farmer a "trained" farmer if he attended a certain session, and "untrained" if he didn't. There are four possible scenarios for each best practice and each session; either a farmer:

- attends a session and adopts a certain best practice (trained and adopted);
- attends and doesn't adopt (trained and didn't adopt);
- doesn't attend and adopts (not trained and adopted);
- doesn't attend and doesn't adopt (not trained and didn't adopt).

If the program is indeed having an impact on farmers we would expect the adoption rate of trained farmers to be higher than the adoption rate of untrained farmers. Figure 15 represents the difference in adoption of trained and untrained farmers over time. We find that while the difference in the adoption rates of trained and untrained farmers was identical at the start of the program, it increased steadily over time.

Another interesting question to ask is: does attending one extra session make a big difference? The answer depends on how TechnoServe defines the objectives of the project. As can be seen in figure 16, if the program objectives were to ensure that farmers adopt at least 5 best practices out of the 11, then the answer is not really. While the marginal increase in the likelihood of adopting at least 5 best practices is always positive, it decreases as the number of sessions attended increase. However, every extra session matters if the objective of the program is to ensure that farmers adopt all 11 best practices. Attending one more session when you have already attended 15, makes much more of a difference than when you move from 2 to 3 sessions for example. The most likely explanation as to why this might be the case is that every session focuses on a different best practice (even though sessions tend to touch on multiple best practices at a time as we will see in the next section); the closer a farmer gets to attending all classes the increasingly more likely it becomes that he/she will be become an adopter of all 11 best practices.

Figure 16: Estimated marginal increase in the likelihood of adopting best practices based on attendance levels



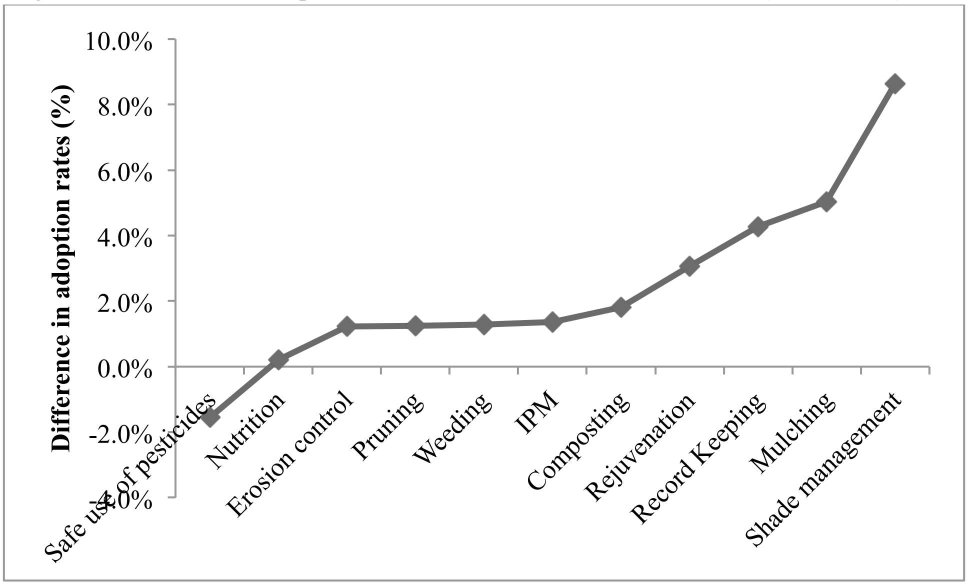
We present strong and statistically significant evidence that there is a link between attendance and adoption. One could argue however that these differences appear to be quite small: a 6 percentage point difference in adoption rates between trained farmers and untrained farmers, or a 6 percentage point difference in attendance rates between adopters and non-adopters, does not hint at a very large impact. The main reason we observe such small differences is that the best practice sample is homogenous and consists of farmers with high attendance levels (farmers in the best practice sample were selected amongst farmers with more than 50% attendance in year 1) and high adoption rates (the average adoption rate was 8.82 best practices out of 11 at the end of the training); if the sample had also included low attendance farmers we would in all likelihood have been able to draw a starker contrast between farmers who attended training sessions and farmers who didn't. Another reason is that we have been focusing so far on aggregate adoption rates regardless of the best practice and the corresponding training session. In the next section we delve into a bit more detail, by ultimately matching specific training sessions to best practice adoption.

(ii) Link between attendance rates and specific best practices

Has TechnoServe's coffee agronomy training program been equally successful in improving adoption rates for all best practices? We focus again on Cohort 2010 for which we have final best practice adoption data, post training (i.e. round 2, 2011).

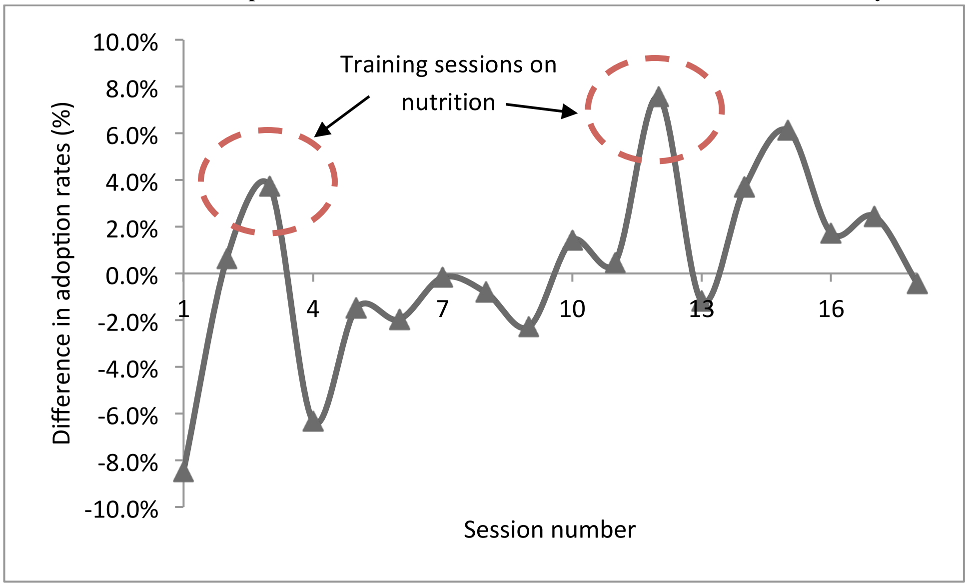
At a first glance, when comparing the average adoption rates of trained and untrained farmers – where we call a "trained" farmer a farmer that has attended a certain training session and an "untrained" farmer a farmer that did not attend that specific training session - we find that the program seems to have been more successful in increasing adoption for some best practices than for others (see figure 17). The average difference in adoption rates between trained farmers and untrained farmers is the largest for shade management, mulching, and record keeping. While we cannot causally attribute this difference to the training, the fact that only 12% of farmers claim to have known about shade management before the program and only 6% about record keeping, suggests that the training did play an important role in increasing adoption rates for these best practices. The best practices with the lowest average differences in adoption rates are safe use of pesticides and nutrition. This could mean that the program was less effective on these best practices or that the adoption of these best practices was hindered by external factors, such as the fact that adoption depends on the acquisition of specific inputs (personal protection equipment in the case of safe use of pesticides, and fertilizers in the case of nutrition). However, we show that it is necessary to go to a more granular level and compare attendance in specific sessions where a certain best practice was taught to fully analyze the link between best practice adoption and training.

Figure 17: Difference in adoption rates of trained and untrained farmers (Cohort 2010)



Zooming-in on how the adoption rates for trained farmers and untrained farmers vary on a given best practice by session (in this case we take the specific training session into account), we find that - in most cases - farmers that attended a session in which the best practice was taught are more likely to adopt that best practice than farmers that did not attend that specific session. The case of nutrition is interesting. While we do not find large differences in the nutrition adoption rates of trained and untrained farmers on average, we find that farmers that attended sessions 3 and 12 – where nutrition was taught – had higher nutrition adoption rates than farmers that did not attend those sessions. As can be seen in figure 18, the difference in nutrition adoption rates between farmers that attended a session (i.e. trained farmers) and farmers that didn't (i.e. untrained farmers), spikes in sessions 3 and 12. This strongly suggests that nutrition training had a lasting impact on nutrition adoption in Cohort 2010.

Figure 18: Difference in Nutrition adoption rates between trained farmers and untrained farmers by session (Cohort 2010)



Controlling for the effect of time, we calculate the difference in adoption rates of trained and untrained famers for each best practice in each session in Cohort 2010. The two sessions in year 1 of training and the single session in year 2 of training where the differences in adoption rates between farmers that attended a training session and farmers that did not was the largest are highlighted in orange (figure 19). Think of these as the training sessions that were the most effective in increasing adoption rates for a given best practice. We find that the most effective training sessions for many best practices are also the ones corresponding to the best practice at hand: attending the IPM training session for example is associated with higher IPM adoption rates; attending the nutrition session is associated with higher nutrition adoption rates; attending the weeding session is associated with higher weeding adoption rates; attending the composting session is associated with higher composting rates, etc.

Figure 19: Matrix of most effective training sessions based on theory vs. in practice

Best Practice > / Training sessions v	Record Keeping	Mulching	Weeding	Nutrition	Composting	Rejuvenation	Pruning	Safe use of Pesticides	IPM	Erosion Control	Shade Management
Record Keeping (Y1)	X										
IPM (Y1)	X			X		X	X	X	X		X
Nutrition (Y1)	X	X	X	X	X				X		X
Harvesting (Y1)	X										
Weeding (Y1)		X	X	X							
Mulching (Y1)		X	X	X						X	
Rejuvenation & Pruning (Y1)						X	X		X		
Safe use of pesticide (Y1)	X							X			
Composting (Y1)				X	X				X		
Erosion Control (Y1)		X								X	
Shade management (Y1)			X	X					X	X	X
Harvesting & Nutrition (Y2)	X	X	X	X	X				X		X
Harvesting & Nutrition (Y2)	X	X	X	X	X				X		X
Sustainability (Y2)											
Sustainability (Y2)											
Rejuvenation & Pruning (Y2)						X	X		X		
Lime (Y2)											
Profit & loss (Y2)	X										

Key: [orange] Training sessions that were most effective in increasing adoption rates for a given best practice
X Training session that are most effective in increasing adoption rates for a given best practice in theory

More than just the link between a specific session and a best practice, the matrix above also highlights the inter-linkages between best practices. As figure 19 reveals, what we observe in practice is also closely linked to theory. We find that the training session that are most effective in increasing adoption rates for a given best practice based on agronomic theory (i.e. denoted by "X" in the matrix) closely relate to our findings based on actual data (i.e. denoted by the orange highlight in the matrix).

Did You Know?*

- **Did you know that the main BPs that farmers perceive they acquired from TechnoServe trainings are Composting, Safe Use of Pesticides, Pruning and Mulching?**

Best Practice	Main best practice acquired from training (multiple answers permitted)
BP 1: Record keeping	14%
BP 2: Mulching	**27%**
BP 3: Weeding	3%
BP 4: Nutrition	22%
BP 5: Composting	**30%**
BP 6: Rejuvenation	9%
BP 7: Pruning	**28%**
BP 8: IPM	24%
BP 9: Safe use of Pesticides	**30%**
BP 10: Erosion Control	6%
BP 11: Shade Management	15%

- **Did you know that farmers considered Mulching and Nutrition as the most important BPs for increasing yield levels?**

Best Practice	Most important BP to increase yields (multiple answers permitted)
BP 1: Record keeping	2%
BP 2: Mulching	**55%**
BP 3: Weeding	17%
BP 4: Nutrition	**34%**
BP 5: Composting	13%
BP 6: Rejuvenation	6%
BP 7: Pruning	22%
BP 8: IPM	8%
BP 9: Safe use of Pesticides	8%
BP 10: Erosion Control	2%
BP 11: Shade Management	1%

- **Did you know that farmers considered Mulching to be the most difficult BP to implement?**

Best Practice	Most difficult BP to implement?
BP 1: Record keeping	15%
BP 2: Mulching	**37%**
BP 3: Weeding	1%
BP 4: Nutrition	10%
BP 5: Composting	7%
BP 6: Rejuvenation	2%
BP 7: Pruning	8%
BP 8: IPM	7%
BP 9: Safe use of Pesticides	7%
BP 10: Erosion Control	2%
BP 11: Shade Management	5%

*Based on spot-check analysis data

Chapter 6: The Monitoring Effect

CHAPTER SUMMARY:

One of the most unexpected and surprising results of this study is the large effect of monitoring on the way farmers experience the training program. TechnoServe's extensive and unique monitoring and evaluation system in Rwanda enables us experimentally conduct one test that even the best randomized control trials would not allow: to test the impact of the monitoring and evaluation system itself on project beneficiaries. We find that monitoring on farmers in the yield sample - characterized by frequent interaction between farmers and TechnoServe staff - has a significant impact on how these farmers experience the training. Monitoring on yield farmers leads to a 12-percentage point increase in attendance rates and a 7-percentage point increase in best practice adoption. This has practical implications on fertilizer usage too, as we find evidence that monitoring also leads to large increases of NPK, lime and Zn/Bo usage. Conversely, less frequent monitoring on farmers in the best practice sample, also leads to increases in attendance. The findings on the impact of monitoring raise interesting questions about how TechnoServe can leverage the monitoring effect to improve project outcomes.

TechnoServe's extensive and unique monitoring and evaluation system in Rwanda enables us to conduct one test that even the best randomized control trials would not allow: to test the impact of the monitoring and evaluation system itself on project beneficiaries. TechnoServe has diligently collected extremely detailed attendance data on the entire "population" of registered coffee farmers. For each coffee farmer in the project, we know which of the 16 to 18 training sessions they attended and whether they attended a given session alone or together with their spouse. This alone amounts to more than 870,000 data points on attendance for Cohorts 2009, 2010 and 2011 (TechnoServe is currently collecting its 1,000,000th data point on attendance with data from Cohort 2012)! We exploit this data to test the difference in attendance patterns for farmers that were randomly selected into either the "yields sample" or the "best practice sample". We find that belonging to the yields sample increases average attendance rates by an estimated 12-15 percentage points, whereas best practice sample farmers exhibit small and short-lived increases in attendance – the subsequent analysis will show that these high attendance rates is due to the monitoring of project beneficiaries by either farmer trainers and/or enumerators. These findings have interesting implications for project design and raise the question of whether the strength of the monitoring effect is inherent to the Rwandan context or whether it also applies to other countries where TechnoServe operates.

(i) Impact of monitoring on the yields sample (on attendance)

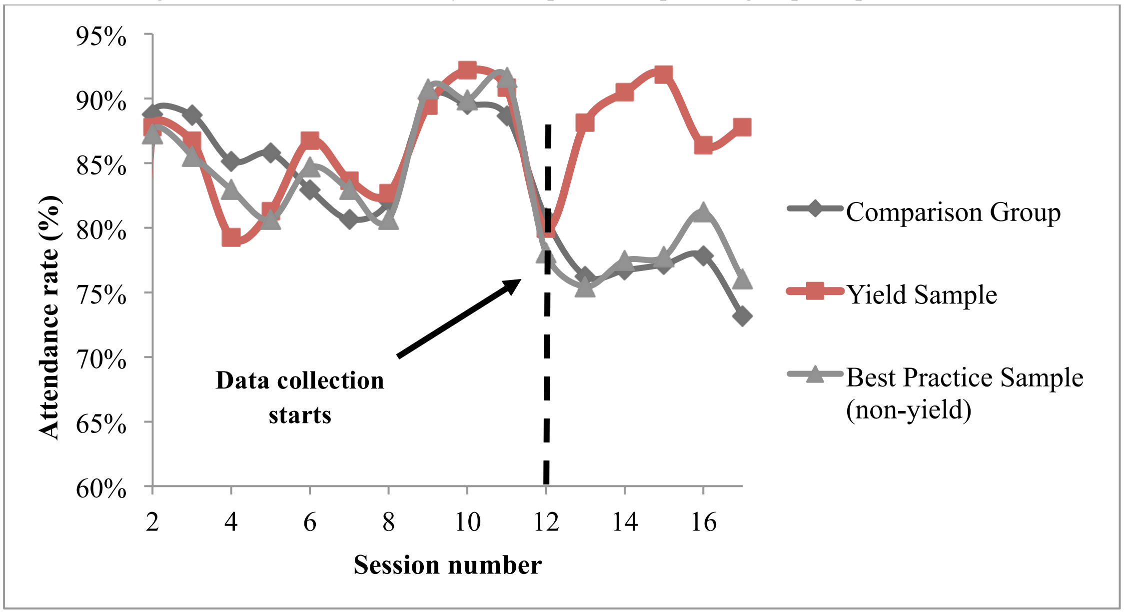
We start by analyzing the impact of monitoring on attendance in the yields sample in Cohort 2009. In Cohort 2009, farmers were assigned to the yields sample after 1 year of training and were selected amongst farmers with attendance rates of more than 50% combined with best practice adoption rates of more than 50%. First, farmers with more than 50% attendance were randomly selected into the best practice sample in each participating cooperative; randomization was done at the cooperative level. Then, after best practice data was collected, farmers with adoption rates of more than 50% were randomly assigned to the yields sample. While the selection process was not fully random, as it resulted in an over-representation of cooperative members and fewer women, this set-up nevertheless enables us to test the impact of belonging to the yields sample on attendance rates.

The reason monitoring on the yields sample is susceptible to changing the way a farmer experiences the project is because of his/her repeated interaction with the farmer trainer over a long period of time and the somewhat more formalized nature of the farmer's relationship to the project. Farmers that are assigned to the yields sample are asked to sign a "scales contract" committing them to keeping records for three years as per TechnoServe's instructions, in exchange for a weighing scale. TechnoServe staff then gives the farmer a weighing scale as well as a paper calendar where the farmer can record the cherry harvested on each day of that month. At the end of each month during the harvesting season TechnoServe staff re-visits that farmer, collects the calendar and gives a new one for the coming month. A TechnoServe Farmer Trainer also checks in on the farmer once in the middle of the month to ensure they are recording the yields properly. Once during the season TechnoServe also does a full tree count, physically counting all of the productive trees on the farm.

If, in addition to attendance data, we had best practice information on all registered farmers in Cohort 2009, it would have been possible to create an almost perfect control group for the yield sample, by selecting farmers with an attendance rate of more than 50% and a best practice adoption rate of more than 50%. In order to correct for the fact that we do not have best practice information on the entire population, we raise the attendance barrier slightly and compare farmers in the yield sample to other farmers with an attendance rate of at least 60% - this will be our comparison group. As we established in the previous chapter, attendance is positively associated with best practice adoption; hence a sample consisting of farmers with an attendance rate of more than 60% should on average have higher best practice adoption rates than a sample consisting of farmers with an attendance rate of more than 50%. According to our estimates, there is a 92% chance that farmers with a minimum attendance of more than 60% would have an adoption rate of at least 50%.

With the comparison sample in place, we conduct two tests: we first compare the evolution of attendance in the "yield sample" over time to that of the comparison group, before conducting a placebo test and comparing the attendance of "non-yield sample" farmers in the best practice sample (with an attendance rate of more than 60%) to the comparison group.

Figure 20: Attendance rates in yield sample vs. comparison group and placebo

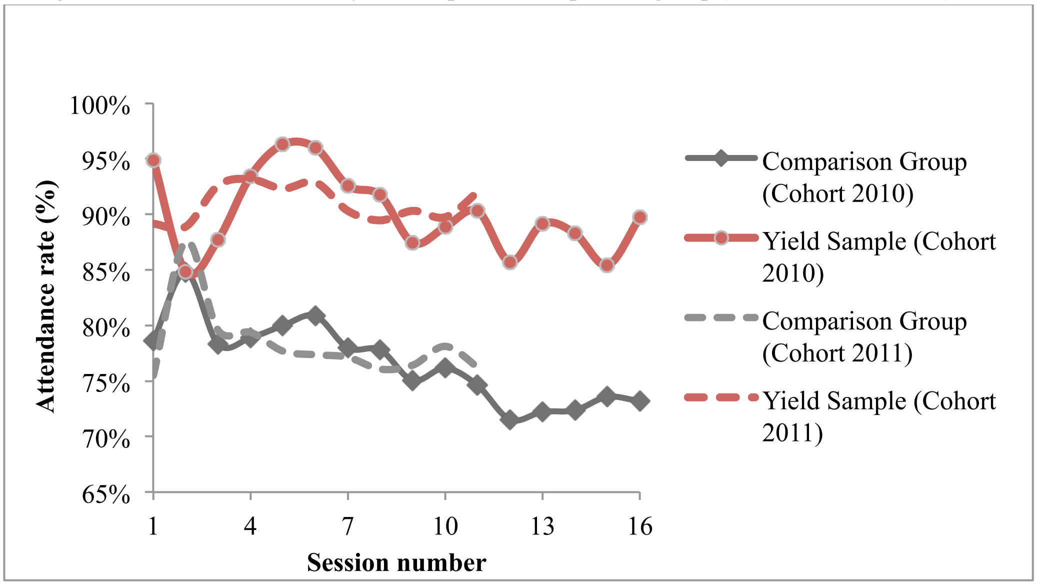


The results suggest that belonging to the yields sample leads to a 12-14 percentage point increase in attendance. While we cannot distinguish attendance rates between farmers in the yields sample, the comparison group and the non-yield best practice sample before yield data collection started, we see a very clear jump in attendance for farmers in the yields sample after the start of the data collection effort (see figure 20). The fact that this is not fortuitous is confirmed by the placebo test, comparing non-yield farmers in the best practice sample (from which the yields sample was derived) to farmers in the comparison group. We interpret this difference as the impact of monitoring and, in particular, the frequent interaction between farmers and TechnoServe staff related to the farmer's involvement in the project.

We confirm these findings with data from Cohorts 2010 and 2011. The case of Cohorts 2010 and 2011 is somewhat different in that farmers that registered in the program were randomly assigned to the yield sample at the beginning of the project, not one year after. In both Cohorts 2010 and 2011, farmers in the yield sample were randomly selected amongst farmers that attended session 1, which is held in January of the corresponding Cohort year[12]. Given that not all registered farmers attend session 1, farmers in the yield sample will have higher session 1 attendance rates than farmers in the rest of the population. While farmers in the yield sample were selected amongst farmers that attended the January session, yield data collection however did not start until March - at the beginning of the harvesting season - which is when TechnoServe typically runs session 3. If the findings from Cohort 2009 hold, we should observe higher attendance rates for farmers in the yield sample starting from session 3. This means that there is only one point in time – i.e. in February, which corresponds to session 2 – where we can compare farmers in the yield sample to other famers without the effect of monitoring.

The problem is that by virtue of the fact that farmers in the yield sample attended session 1, they will have higher attendance levels than farmers outside the yield sample. In order to adjust for this imbalance and create a valid comparison group, we compare the attendance rates of farmers in the yield sample to all other farmers that attended either session 1 or session 2.

Figure 21: Attendance rates in yield samples vs. comparison group (Cohorts 2010 & 2011)



As can be seen in figure 21, the results are very much in line with expectations. While it is not possible to distinguish between the attendance rates of farmers in the yield and comparison groups in session 2, the attendance gap between the two groups increases significantly as soon as the data collection efforts starts and remains constant throughout the program. The average difference in attendance between farmers in the yield sample and farmers in the comparison group is about 13 to 15 percentage points, almost equivalent to the case of Cohort 2009.

We conclude that the type of monitoring associated with yield data collection – i.e. regular, structured and agreed upon in writing – has a significant impact on the way farmers experience the project. This 12 to 15 percentage point increase in attendance, induced by monitoring, implies that the yield impact estimates obtained in chapter 2 are not representative of the impact of the program on the average farmer: they are representative of the impact of the program augmented by a regular monitoring regime. With the available data

[12] In some cases farmers that did not attend session 1, but attended session 2 were included, but this constitutes a minority.

it is unfortunately not possible to provide a breakdown of what share of the estimated impact on yields is due to the training itself and what share can be attributed to the monitoring effect.

(ii) Impact of monitoring the "yields sample" (on best practice adoption)

To test whether the monitoring effect on yield sample farmers has had an impact on best practice adoption or not, we compare best practice adoption rates for farmers in Cohort 2010 to farmers outside the yields sample. By doing so we are comparing a random group of farmers selected at the very start of the program – i.e. yield sample farmers - to a random group of farmers that attended at least 50% of training sessions in year 1 – i.e. best practice sample farmers. All other things equal, we would expect farmers in the best practice sample to have higher best practice adoption rates than farmers in the yields sample, given that they were selected amongst farmers with an attendance rate in year 1 of at least 50%. Yet as we have just seen, monitoring has a very large effect on attendance. Controlling for individual, project related, and cooperative level characteristics, we find that the average attendance rate of farmers in the yield sample was 7.6 percentage points higher than farmers in the best practice sample. Did this attendance advantage combined with the benefits of more direct and regular interaction with TechnoServe staff also translate into higher adoption rates?

A simple regression, controlling for individual, project related, and cooperative level characteristics, reveals that being in the yield sample is associated with an increase in best practice adoption of 0.6 out of 11 best practices, or an increase of about 7% compared to farmers outside the yields sample. While the difference in adoption is statistically significant, when we break down these differences by best practice we find that the yield sample only performs better on nutrition, composting, safe use of pesticides and shade management (see table 24). Nevertheless, given this is likely to be an underestimation of the effect of being in the yield sample on best practice adoption, as we are comparing them against high attendance farmers selected after 1 year of training. See Annex B for detailed regression results.

Table 24: BP Adoption rates (Yield vs. Non-Yield Sample)

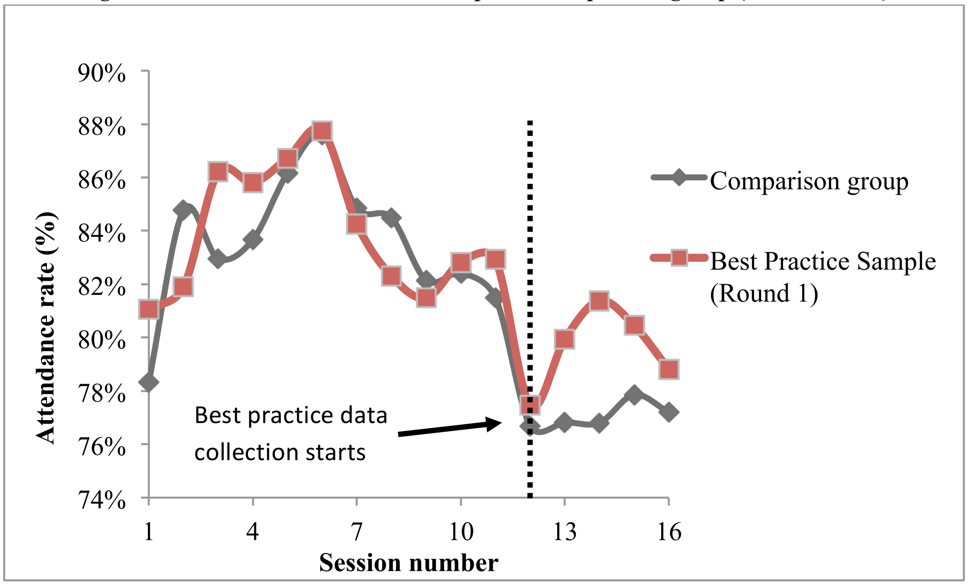
Best Practice	Yield Sample	Non-Yield Sample	Difference
Record Keeping	76.6%	77.4%	- 0.80%
Mulching	92.9%	90.7%	+2.20%
Weeding	98.5%	97.1%	+1.40%
Nutrition	87.3%	82.2%	+5.10%
Composting	68.5%	65.5%	+3.00%
Rejuvenation	89.3%	88.5%	+0.80%
Pruning	84.3%	84.6%	- 0.30%
Safe use of pesticide	77.7%	77.5%	+ 0.20%
IPM	78.2%	74.8%	+ 3.40%
Erosion Control	97.5%	96.9%	+ 0.60%
Shade Management	59.4%	47.4%	+ 12.00%

Another significant difference between yield sample farmers and farmers outside the yield sample – in particular in Cohort 2010 - is that they use significantly more fertilizer, including NPK, lime and Zn/Bo. On average, 50% of farmers in the Cohort 2010 yield sample used NPK, compared to just 32% of other farmers in the best practice sample; 37% of yield sample farmers applied lime, compared to just 14% of non-yield farmers; and 17% of yield sample farmers applied Zn/Bo, compared to 9% of other best practice farmers. While we find a similar patterns in Cohort 2011, the differences between yield farmers and non-yield farmers is much smaller. These findings suggest that the repeated interaction between project staff and yield farmers has a highly positive effect on the likelihood of the latter purchasing and applying fertilizers. This difference in fertilizer usage also suggests that the monitoring effect also leads to large increases in yields, although the scale of this potential increase is difficult to isolate.

(ii) Impact of monitoring on the best practice samples

To test whether monitoring also has an impact on farmers in the best practice sample, we compare the attendance rates of farmers in Cohort 2010 for whom best practice data was collected from March to June 2011 to a comparison group of farmers for whom best practice data was never collected. Farmers in the best practice sample were randomly selected amongst farmers that had attended more than 50% of sessions in year 1 of training; the appropriate comparison group therefore consists of farmers outside the best practice sample but with similar attendance levels.

Figure 22: Attendance rates in BP sample vs. comparison group (Cohorts 2010)



We find that the impact of monitoring on the attendance rates of farmers in the best practice sample is smaller and short-lived compared to the impact of monitoring on yield farmers (see Figure 22). The increase in attendance is limited to about 5 percentage points and seems to fade away after a few months. This is not surprising given that best practice monitoring is significantly less intense than yield data collection: best practice data is collected once to twice per year and on a much larger sample of farmers, thereby limiting the interaction between the farmer and TechnoServe staff. Nevertheless, even the low intensity of this monitoring leads to a significant increase in attendance.

Anecdotal evidence from the focus groups and in-depth interviews confirm the fact that monitoring has an impact on how the farmers experience the program.

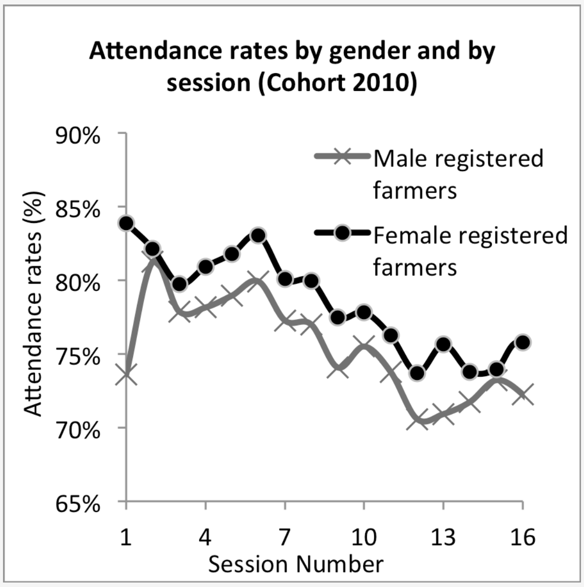
- *"When the trainer comes to me, I know that he will ask me the record book. It gives me a wake up call to check on my farm"*
- *"The presence of the trainer is a reminder. His presence is helpful"*
- *"I always worry if the trainer has noticed that I have failed compared to other farmers in the group, so I work hard to be proud"*

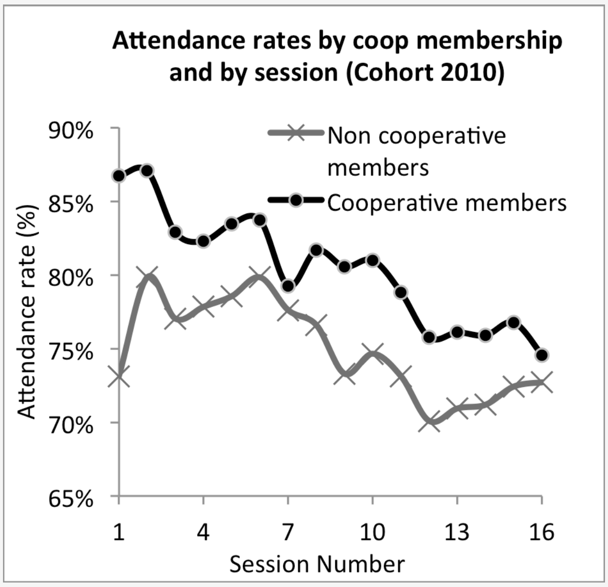
- *"I am ashamed when the trainer sees anything wrong with my BP adoption. I am so happy when I get down to my farm and see that it looks all green!"*
- *"If the trainer comes for evaluation every day, it would make my work harder because it would be a wake up call!"*

The question for TechnoServe moving forward is how to leverage the potential of this monitoring effect to increase project outcomes? We present some thoughts on how this could be done in the conclusion.

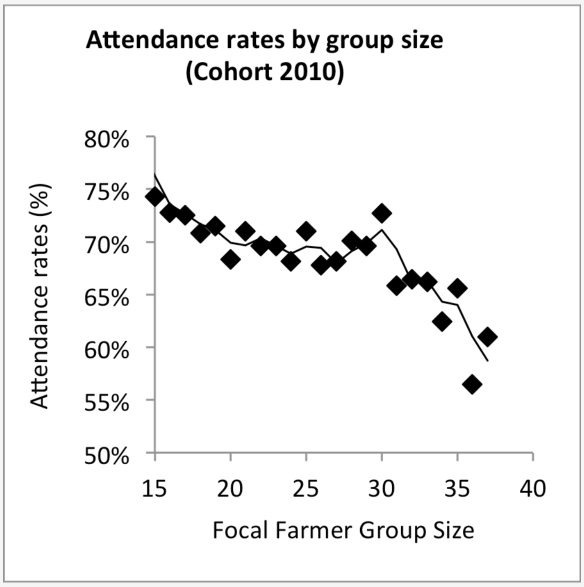
Did You Know?

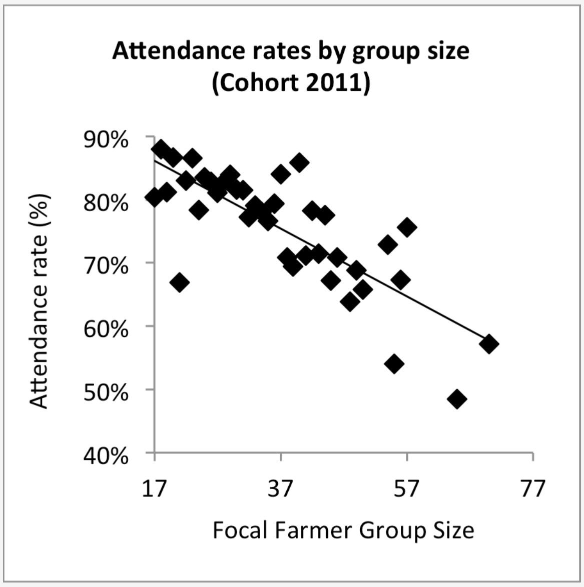
- **Did you know that enthusiasm for training tends to decrease over time, that female registered farmers consistently have higher attendance rates than male farmers, and that cooperative members are more likely to attend than non-cooperative members**?





- **Did you know that group size matters a lot?** The larger the training group, the lower attendance is on average. This might simply mean that in areas where there are larger groups many more people initially sign up to participate in the program but then don't turn up, resulting in lower attendance rates; but it could also mean that the training is less effective in larger groups and that therefore people eventually drop out.





Chapter 7: Conclusion & Recommendations

Summary of results

We find convincing evidence that TechnoServe's agronomy training program has had an impact on agronomy practices and yields in participating cooperatives. Yield levels for a sub-sample of comparable farmers increased by 57.5% in Cohort 2010 and 75.5% in Cohort 2011 during the first year of TechnoServe training. The shift in the distribution of yields from baseline levels to post-training levels was remarkably similar for both Cohorts 2010 and 2011, suggesting that successive project Cohorts are comparable and that the delivery of the training has been consistent over time. However, we do not find evidence of a significant impact of the program on yields in year 2 of training program, although year 1 yield achievements were maintained (this assessment is based on Cohort 2010 alone). An innovative technique based on the similarity of best practice adoption patterns between pairs of farmers also shows that yield increases are intricately related to best practice adoption. While the results are clearly indicative of a causal impact, it is important to note that these impact estimates might be slightly over-stating the actual impact of the program, as it is not possible to distinguish between the effect of the training program and the additional effect of monitoring.

Reinforcing this link between training and yields is the fact that we find a very robust association between attendance and best practice adoption in all project Cohorts. In particular, we find that: (i) the higher a farmer's attendance rate, the greater the likelihood that he/she will adopt a best practice; and (ii) the adoption rate of trained farmers on a specific best practice, is higher than the adoption rate of non-trained farmers. Furthermore, we find increasing returns to attendance when looking at the likelihood of a farmer adopting at least 9, 10 or 11 best practices. Given that the training calendar is structured around different best practices in each training session, these increasing returns are further evidence that the program is strongly associated to an increase in best practice adoption. These data-driven findings are substantiated by anecdotal evidence collected through spot-checks, interviews and focus groups in the field. For all best practices, more than 50% of farmers claimed to have acquired the latter through TechnoServe's agronomy training program. In the case of the relation between training and best practice adoption, it is important to note that there is a high possibility of spill-over effects, running from participant farmers to non-participants. Any estimates therefore underestimate the actual effect of the program.

We do not find conclusive evidence on farmer trainer over-reporting in either the yield or best practice samples, but do find a very large impact of project monitoring on the way that farmers experience the training. Specifically, monitoring on farmers in the yield sample – which consists of bi-monthly (twice per month) interactions between the farmer and project staff over a sustained period of time- led to a 12 to 15 percentage point increase in farmer attendance rates, a 7 percentage point increase in best practice adoption for these farmers and a large increase in fertilizer usage. In the same way, monitoring on farmers in the best practice sample led to a temporary 5 percentage point increase in their attendance rates. Monitoring on farmers in the best practice sample, however, only happened twice per year. These findings lead to interesting questions about how the project can leverage these monitoring effects to improve project outcomes, but also raise the question of whether this same monitoring effect applies at other levels in the program structure – e.g. the focal farmer level and more importantly the farmer trainer level.

Proposed changes and improvements

Monitoring & Evaluation Recommendations

- ***Collecting data on basic individual farmer characteristics***. In trying to argue the case that TechnoServe's Cohorts are good control regions for each other and in breaking down the analysis to the individual level, one of the main pieces of information our team was lacking were simple additional (and easy to collect) individual farmer characteristics such as age, distance from home to the training plot, family status, main sources of income, experience in growing coffee, etc. We recommend that TechnoServe collect such data for each new batch of participants, either at registration or at the beginning of the project when the best practice sample is formed. This type of information will increase TechnoServe ability to argue that its monitoring and evaluation approach – where one Cohort is used as a control for the next – is valid. Moreover, more information on who their participants are will increase TechnoServe's understanding of best practice and yield dynamics at the individual level and enable the organization to better tailor and test its training methods.

- ***Adjustments to how best practice data is collected***. We recommend four adjustments to the way TechnoServe currently collects best practice data:

    - ***Selection of farmers: there is no need to focus on farmers with attendance levels of more than 50% in year 1.*** To date - Cohort 2012 aside - best practice data has been collected from farmers that attended at least 50% of training sessions. While this does not lead to major biases, focusing solely on farmers with high adoption rates has several disadvantages: (i) the resulting sample is not representative of the entire training population; (ii) it limits our knowledge of what is happening for farmers with lower attendance rates; (iii) it leads to lower variation amongst farmers in the best practice sample, thereby limiting the ability of an analyst to study the impact of the project or other related dynamics; and (iv) it is unnecessary, because the vast majority of farmers have more than 50% attendance levels in year 1 anyway.

    - ***Baseline data: collecting baseline best practice information on farmers in the best practice sample will provide necessary context.*** While TechnoServe does conduct baseline surveys on agronomic practices in participating cooperatives, these surveys are conducted before farmers have self-selected into the program and use slightly different metrics to the metrics used in the best practice sample. This makes them difficult to compare to the post-training best practice data that is collected only one year after the start of the program. Having baseline best practice data on farmers in the best practice sample will enable researchers to study the impact of the project and related dynamics in much more detail.

    - ***Timing: TechnoServe can reduce monitoring costs by concentrating its best practice data collection efforts on the months of March and July, or by just collecting best practice data once every year.*** While TechnoServe's best practice data collection effort is remarkable, the fact that best practice data is collected almost constantly in the field (from March through to June, and then again in July, August, September and November) leads to comparability issues. Best practice adoption rates fluctuate quite a lot from month to month, which means that even if best practice data is collected in the same "round", the timing of the data collection can somewhat skew results. Best practice adoption peaks in March and July, which means that these are probably the best months of the year to collect best practice data. Moreover, while it theoretically makes sense for TechnoServe to collect data twice per year in order to capture seasonal differences, it is not clear what the practical benefits derived from this data collection effort are. If the main purpose of this data collection effort is to report on average best practice adoption rates, then we would recommend reducing the monitoring burden and collected data only once per year, either in March or July.

- ***Metrics: we strongly recommend that TechnoServe tweak its current metrics and ideally also make its checklist more specific, thereby reducing the risk of human error but also increasing variation in the best practice sample which can lead to interesting and useful information for the project.*** We recommend several changes to the best practice Monitoring & Evaluation (M&E) process namely (i) adjusting metrics in the best practice checklist to eliminate current inconsistencies; and (ii) providing more granularity in the best practice check-list. On point 1 there are currently inconsistencies in the way best practices 6, 8, 10 and 11 are measured (see table 25). The problem is that farmers for whom a given best practice does not apply (e.g. farmers that are not on a slope do not need to worry about erosion control) are nevertheless counted as adopters, which clouds actual adoption rates. On point 2, more granularity or much higher adoption standards (a "gold standard" of adoption) will provide TechnoServe with more variation in farmer adoption rates. TechnoServe can use this variation to better understand dynamics in the field and compare the performance of the project across cooperatives on more specific metrics.

Table 25: Suggested changes to Best Practices Adoption metrics

Best Practice	Options	Current Metrics	Suggested Metrics
BP 6: Rejuvenation	No rejuvenation, trees with main stems of 8 years or older	Did not Adopt	Did not Adopt
	Trees Rejuvenated in the last 8 years	Adopted	Adopted
	Trees recently planted, not ready for rejuvenation	Adopted	N/A – Drop from analysis
BP 8: Safe Use of Pesticides	Pesticides used, no PPE used or containers not disposed of correctly.	Did not Adopt	Did not Adopt
	Pesticides used with PPE, containers disposed of correctly	Adopted	Adopted
	No Pesticides used	Adopted	N/A – Drop from analysis
BP 10: Erosion Control	No erosion control methods being used	Did not Adopt	Did not Adopt
	Erosion control not required, land flat	Adopted	Adopted
	Land not flat, at least 1 method of erosion control used: terraces, grasses, mulch, water traps, Physical barriers e.g. rocks	Adopted	N/A – Drop from analysis
BP 11: Shade Management	No shade trees or heavy shade over 40%	Did not Adopt	Did not Adopt
	10-20% shade or shade trees planted recently	Adopted	Did not Adopt
	20-40% shade, established trees (include bananas)	Adopted	Adopted

- ***Cost-permitting, increase the sample size of the yield sample.*** We are aware of the cost implications of increasing the sample size of the data collection effort on yields, but depending on what type of analysis TechnoServe expects from this data, this might be useful. The problem with the current yields dataset is that it consists of about 30 farmers per cooperative, with very high levels of variation within them. Yield levels typically vary from 0.2kgs/tree all the way through to 6-7kgs/tree in every single cooperative. This means that it is very difficult to conduct any type of analysis on yields at the cooperative level. Moreover, the level of variation combined with the small sample size makes it difficult to obtain statistical significance, for example of the relation between attendance and yield levels.

Leveraging the monitoring effect

Is it possible for TechnoServe to obtain the benefits of monitoring without incurring the costs? Given the impact of regular yields and best practice monitoring on attendance and best practice adoption rates, creating a placebo monitoring effect might be an interesting tool for TechnoServe to explore and test. As shown in the chapter on the monitoring effect, current field visits by TechnoServe staff act as a reminder and motivation for farmers to attend training and also ensure that they are implementing the best practices on their farms. Bearing in mind the high costs of direct monitoring, we propose conducting further research to test alternatives that could potentially result in an "illusion" of monitoring and lead to similar outcomes.

Annex A: Focus Group & Semi-Structured Interview Transcripts

As part of the spot-check process, we conducted one focus group and 2 semi-structured interviews in each of the selected cooperatives. The objective of these additional research tools was to ask specific questions to find potential explanations as to why we observe certain discrepancies/outliers in the data. Each focus group consisted of between 5-7 randomly selected farmers to determine the reasons behind best practice adoption and non-adoption, interactions with farmer trainers, understand perceptions of the TechnoServe program, learning and attendance dynamics. We have compiled the results of these interviews into the following components:

Cooperative: CAFEKI

Component	Details	Quotes
1. Coffee Farming Knowledge	Where did you acquire the best practice knowledge from?	*"Mainly from TechnoServe and few things from our parents, friends or neighbors"*
	Learn on your own?	*"Before we got trainings, we knew how to mulch, fertilize, weeding, pruning and harvesting"*
	Learn from friends/neighbors?	*"We generally got motivation from neighbours to become coffee farmers: rejuvenation, pruning, erosion control and cleaning is what we learnt from our neighbours"*
	Learn from TechnoServe?	*"We knew how to do some of the BPs, TNS showed us how to do it better and make sure it worked"*
2. Assessment of BP adoption and non-adoption	Which BPs have low adoption?	**Composting**: *We do not use composting because of our mindset - we all think that composting is only what comes from the cows but we never learnt techniques of how to make other types of composting. The second reason is that most of the times we have to buy it and it is very expensive to buy it for those who don't possess cows"* **Composting/Fertilizer**: *we don't use fertilizer or manure because it is expensive to buy it; also the local authorities provide it on loan but the way they do it is not ideal for farmers because if you don't make the payment time they take your coffee by force so some farmers fear the loans – there is no payment installation scheme too"* **Record Keeping**: *Lastly we don't keep records because farmers don't know exactly why they have been given these books - they don't value it. We used to write in the book only on training days. Another reason is that some of them are not able to read and write* *We think that the trainings on this were not sufficient - we would like to invite TNS to train us again especially on records keeping.*
3. Perception of TechnoServe trainings	What do you like most about TNS training?	*"What I liked the most was the pruning techniques in the coffee farm"* *"I liked every single thing we learnt from TNS"*
	What do you like least? Areas for improvement?	*"We need more trainings on how to use fertilizer and manure because no one knows how to use it in our region, so, we need more training"*

Cooperative: COCAMU

Component	Details	Quotes
1. Coffee Farming Knowledge	Where did you acquire the best practice knowledge from?	*"Most BP training came from TNS after 2009. Prior to that, we had officials from the district and NAEB but it was not really training, more like sensitization"*
	Learn on your own?	*"We used a lot of traditional techniques that we learnt from our parents and friends but it's only since TNS came we got professional training"*
	Learn from friends/neighbors?	*"I learnt mulching, weeding and pruning from friends and neighbors"*
	Learn from TechnoServe?	*"Mulching, weeding, how to control pests, composting and so many others"* *"IPM and shade management"* *"I learnt from TechnoServe trainings safe use of pesticides. They also told us the good pests and bad pests in coffee trees"; "What I did not know before TechnoServe was, for example; record keeping, shade management, IPM, composting and safe use of pesticides"*
2. Assessment of BP adoption and non-adoption	Which BPs have low adoption? **Mulching, Pruning, Safe use of pesticides**	*"For mulching; it is not easy to find mulch. That is why sometimes we don't put mulch"* *"For pruning; we do not have materials, I just use my hands. It is difficult to do it with hands"* *"For the safe use of pesticides, it is the same as pruning, we do not have materials. No PPE used"*
3. Perception of TechnoServe trainings	What do you like most about TNS training?	*"TechnoServe came to answer all the questions we used to ask ourselves about coffee farming techniques. But particularly, I liked record book, IPM and composting."* *"I liked the way we used to meet up in someone's field and have training there. TechnoServe trainers trained us by showing us how to practice on the field."* *"They showed us how to record all activities concerning coffee farming. They taught us pruning, composting, IPM and so many other best techniques."*
	What do you like least? Areas for improvement?	*"Some farmers live very far from where trainings took place. Sometimes some farmers didn't attend the trainings because of that."* *"Presently we don't see TNS trainers, so we don't know whether trainings have ended or not. Another issue is that some farmers don't attend the trainings as they live very far from where the trainings took place - so we request TNS to hire more trainers so they can reach more coffee farmers."* *"We also ask TNS to give us materials in order to adopt some best practices; like pruning, IPM."* *"We ask TNS to give us materials so that we can adopt these best practices. For example, I cannot prune with my hands, it is not easy. I do it but it is not easy."*

Cooperative: COOCAFE

Component	Details	Quotes
1. Coffee Farming Knowledge	Where did you acquire the best practice knowledge from?	*"Between 2007/8, attended training provided by OCIR and Stabex – since then it's only been TechnoServe"* *"Attended TechnoServe training in 2009. Even though it is difficult for me to reach to my coffee farm because I am old - 101 years old!"*
	Learn on your own?	*"Applying compost from cows"* *"Spraying pesticide at the right time; weeding"* *"Killing other pests by using chameleons"* *"Pruning and other traps"*
	Learn from friends/neighbors?	*"Pruning"*
	Learn from TechnoServe?	*"Honestly, most of the best practices I learnt from TechnoServe: composting, IPM, shade management;… there are so many"* *"From pruning to nutrition; killing pests effectively, shade management and record booking"* *"Mulching and shade management"*
2. Assessment of BP adoption and non-adoption	Which BPs have low adoption? **Nutrition, Mulching, IPM**	**Nutrition** – *"the main reason behind the non adoption of this best practice is financial means; It is very expensive for us to buy fertilizers; When you don't own a cow how can you make manure? It is hard to do composting plus NPK is expensive!"* **Mulching**: "*trees are so many compared to the grass we have for mulching. It doesn't even cover half of the farm; Mulching is very crucial to the life of a coffee tree, yet we are failing to do it. It is very expensive and it is also hard to find banana leaves because I adopted coffee farming and left banana farming, although the leaves are very good for mulch; It is very expensive: I can not find two million RWF only for mulching like Rwakajoma did! (Rwakajoma is a focal farmer who spent more than RWF 2M to just buy mulch for his entire coffee farm!"* **IPM**: *"The only reason we don't spray pesticides is because we get them at the wrong time. Simply because pesticides are provided to us when the coffee is already grown - and it is not good to spray a grown coffee tree with pesticide as it could affect its quality"* *"Pesticides come very late!"*
3. Perception of TechnoServe trainings	What do you like most about TNS training?	*"The most important thing I learnt from TNS is mulching"* *"I loved how they teach us; everyone asks questions freely; the trainer is interactive"* *"I loved the way they taught us how to kill pests - I personally didn't know that"* *"I liked the training about how to do record booking"*
	What do you like least? Areas for improvement?	*"We have requested bank credit from TNS so that we can buy mulch but we have not got any response"* *"We really lack the materials e.g. pumps; cutters, etc."* *"Spraying pesticides is not safe without protective materials – maybe TNS is in charge!* *One trainer is not sufficient.*
4. Perception of Farmer Trainer	What is the impact when the farmer trainer comes to check on you?	*"When the trainer comes to me, I know that he will ask me the record book. It gives me a wake up call to check on my farm"* *"The presence of the trainer is a reminder. His presence is helpful"*

		"I always worry if the trainer has noticed that I have failed compared to other farmers in the group, so I work hard to be proud" *"I am ashamed when the trainer sees anything wrong with my BP adoption. I am so happy when I get down to my farm and see that it looks all green!"* *"If the trainer comes for evaluation every day, it would make my work harder because it would be a wake up call!"*

Cooperative: KARAMA

Component	Details	
1. Coffee Farming Knowledge	Where did you acquire the best practice knowledge from?	*"Before MINAGRI gave us the BP trainings of farming and taught us about mulching. This also included planting grass to prevent soil erosion."*
	Learn on your own?	N/A
	Learn from friends/neighbours?	*"Our neighbours do not grow coffee plants so I don't get any advice"*
	Learn from TechnoServe?	*"Acquired the farming best practices from TNS besides that we have not got training from any other institutions"*
2. Assessment of BP adoption and non-adoption	Which BPs have low adoption? **Record keeping, IPM**	*Most of the time I do not keep records due to negligence.* *"I do not mulch because I do not have means financially to buy the mulch"* *"I do not prune because I do not have requirements or tools"*
3. Perception of TechnoServe trainings	What do you like most about TNS training?	*"No problem with TNS trainings and my wish is that they can increase the trainings at least once every month"*
	What do you like least? Areas for improvement?	N/A

Cooperative: KINYAGA

Component	Details	Quotes
1. Coffee Farming Knowledge	Where did you acquire the best practice knowledge from?	*"Long time ago, the agronomy govt. official use to come and just tell us to do this and that but TNS came all the way to the field and showed us how to do it. Now we know how to use the NPK, we know how and when to pick our berries from the coffee tree and now we get more yields from the same tree"*
	Learn on your own?	*"Before TechnoServe came in, our coffee didn't have much value - this was because we didn't really understand how important coffee was and so after TNS came in they showed us how to improve our coffee plantations. We used to harvest in a way that was inappropriate."*
	Learn from friends/neighbours?	*"There are several methods that we learnt from our fathers and neighbours, it just that they didn't pay attention to it and simply planted the coffee trees and left them to grow, but some of the methods included mulching and fighting soil erosion"*
	Learn from TechnoServe?	*"I was trained by TechnoServe on different best practices - some BPs we knew but did not know the importance or the right way to implements such as: pruning, canopy management, IPM, mulching etc."*.
2. Assessment of BP adoption and non-adoption	Which BPs have low adoption? **Record keeping, IPM**	**Record Book** – *"the reason not as many were writing in their record book is because there were some people that did not understand why they should fill in the book; other were scared that this is would be a way to be taxed, i.e. if you write everything, then you will have to pay much more after your harvest. And others simply couldn't read or write.* *"Others would not write because they can't read or write at all and no one was there to help them write in their record books."* *"Another reason is the constant changing of prices which were getting lower – some farmers got lazy and felt less motivated to work on the record books"* *"Some people would advise not to write down everything and were spreading the wrong information to other farmers – I don't know why, maybe jealousy"* **IPM**: "*this one is really hard to do because we have other things to do, and IPM really needs you to be in your plantations most of the time. Furthermore, since they lowered the price for a kilo of coffee beans, people don't care as much now. The problem with IPM is that we don't get as much pesticides as we need - even when we do get it, we don't get the equipment to use it with!"*
3. Perception of TechnoServe trainings	What do you like most about TNS training?	*"I have attended most of the trainings because I found it useful and it has resulted in more yields to our harvest"* *"What I liked most was the way they provided us with the possibility of harvesting, taking the coffee to the factory to be processed and then paying us right away"*
	What do you like least? Areas for improvement?	*"When TNS first came in we were happy with their pricing, it was 300Frw/Kg in 2011 and now it is 170frw/kg in 2012. They told us that the price of coffee had gone down on the world market. They*

		should try to keep up the price, which is what motivates farmers. "Another thing is that we don't get enough fertilizers e.g. NPK so. If they could provide it on a loan basis, we would be happy to pay back after our harvest."

Cooperative: KOBAKANYA

Component	Details	Quotes
1. Coffee Farming Knowledge	Where did you acquire the best practice knowledge from?	*"Back in the day, OCIR Café used to come and provide us with some fertilizers but they would just give it to us and leave without showing us how to use it"*
	Learn on your own?	*"The only things we knew before TNS came in were the basics of coffee planting e.g. mulching. We never considered pruning or soil erosion control"* *"The only techniques we knew were rejuvenation and mulching but it was not good - it was what we knew from parents and neighbours.*
	Learn from friends/neighbours?	*"There are others methods that we had adopted from our friends like mulching, and pruning but we found out that even though our friends had showed us this way they were not doing it properly, so we are grateful that TNS came in to help"*
	Learn from TechnoServe?	*"Some of the best practices that gave us more yields are mulching, nutrition, using lime and NPK 17"* *"Also composting - the main reason we do a lot of compositing is because we think that there will be a time when we won't be able to get the NPK or the Zn/Bn, so we are encouraged to use this"*
2. Assessment of BP adoption and non-adoption	Which BPs have low adoption? **Record keeping, Composting & worm tea**	**Record Book** – *"Not everyone can read and write, and some still don't understand the importance of writing in the record books. Most of them got the wrong message from our neighbours or friends telling them that if they tell the truth, they will probably get taxed higher"* **Worm Tea**: **"***Worm tea has not been distributed to all of us, but I did use it once - we have hope that we might be getting in September"* **Composting***: "The only reason that compositing is not used a lot is because most of the time our coffee farms are far from our homes and you have carry it for a very long distance to the plantations"*
3. Perception of TechnoServe trainings	What do you like most about TNS training?	*"What I liked most in the training is the use of fertilizers and lime, this all led to healthy coffee trees - the problem is that we do not get enough fertilizers. And now that the price/kg has gone down so much people are being discouraged in coffee farming and moving into other crop growing to sustain their income"*
	What do you like least? Areas for	*"What I didn't like much about the training is that they made promises and they did not fulfill them, for example, they agreed to*

	improvement?	*get us some equipment like pumps, bicycles, wheel burrow. If they could work on this, it would help alot"*
4. Perception of Farmer Trainer	What is the impact when the farmer trainer comes to check on you?	*"You are more encouraged by the actions of your fellow farmers; for example, if they do their pruning and you have not, you feel bad as your coffee tree looks the worst - sometimes fellow farmers may say that you have been using your money to drink and not work on your plantations, which leads to a bad reputation"*

Cooperative: MATYAZO

Component	Details	Quotes
1. Coffee Farming Knowledge	Where did you acquire the best practice knowledge from?	*"A few people got training from OCIR Café a long time ago but they were not providing training in the best way as they never reached plantations – they just told us a few things and gave some fertilizers and left it all to us to figure out how to do it or how much we should use"*
	Learn on your own?	*"Our old style of taking care of coffee was poor. There is nothing much that I have learnt on my own"*
	Learn from friends/neighbours?	*"Before we used to copy from our neighbours or the ways that our parents taught us - this is because most of our parents were coffee farmers"*
	Learn from TechnoServe?	*"I learnt a lot about mulching because in the past we didn't know how to do this"* *"We have adopted most of these BPs and we have increased yields. The BPs that have really helped is mulching, pruning and the way they show us how to pick the berries without ruining the remaining part of the branch"*
2. Assessment of BP adoption and non-adoption	Which BPs have low adoption? **Record keeping, Composting & IPM**	*The BPs that we don't put into action are those that need additional equipment. For example, composting - some of our plantations are far from home so it would be better if we have a wheel burrow* Composting*: "Compositing; we did not really understand that composting could work like the NPK but now we know because they taught us. They also taught us how to make it with different compositions* Record Book*: "Some of us can't read or write, and others just feel like they don't have time to do write this down. But I think if more training was given it would be a good thing because people would better understand the importance of keeping record"* IPM: *"We know this method is good but it takes too much time and since coffee is no longer the same price /kg, people just got a bit lazy so they are prefer focusing on the crops that can sustain them"*

		Pesticide; we do use this but the hard part is that we have to borrow the pumps and don't have any proper protection equipment and don't even know where to buy them"
3. Perception of TechnoServe trainings	What do you like most about TNS training?	*"The training was good all the time since what they taught us has generated more yields except that the coffee value has gone down"*
	What do you like least? Areas for improvement?	*"Everything has been good except that we don't have equipment - if we could use scissors, pumps and other items, our harvest would increase and we could pay for these items back in a very short time"*

Farmer Profile: Deogratias

Personal

Deogratias is 72 year old farmer who is a Nasho Cooperative member and has been working in coffee for 36 years. Before moving into coffee farming, he was farming other crops such as beans, sorghum and potatoes. However, given the increased prices in the coffee sector, he switched to coffee production. Deo lives with his wife, one 29year old single son (who helps on the coffee farm) and a grandchild. He has 2 married daughters and 2 other sons in the military.

Income

In addition to coffee farming, Deo grows rice and is part of a rice-cooperative. He hires 6 workers who work mainly on his rice farms. Coffee is now his biggest income earner due to the increase in yields. Before TechnoServe came, the coffee yield was sold a to a rich person who would cheat them on the weight of the coffee yields. TNS provided coffee weighing scales that, in addition to training, have ensured that they increase their yields and are not cheated on their yields. He notes that there is a benefit to being part of the cooperative as they benefit from higher prices as they are organized and have negotiated better rates (but the coop president is not very active and doesn't interact with the focal farmers as he should).

Focal Farmer

Deo is the elected focal farmer for 30 farmers in his sector. As a focal farmer, he states that his role is to (i) visit and encourage farmers to visit his farm (ii) for training, the farmer trainer would call him and he would inform all the other farmers about the training (iii) he also informs farmers of the day that they will apply the pesticides, as pumps are transferred between groups (iv) the farmers will inform him when their coffee is ready to be sold so they can arrange a day to visit the wet-mill together to sell the coffee (iv) resolve any disputes between the farmers and the wet-mill contact.

Knowledge

Before TNS, there were a few government agronomists who would come to see the quality of the coffee tress and would provide some advice but would not provide training. He had some knowledge of coffee farming by learning on his father's farms in the 1950s in northern Rwanda, before moving to the Nasho area. No other NGO provided training in the area.

TechnoServe training

Deo regularly attends the monthly training. He is usually not aware what the topic of the training will be in advance. The main reasons farmers don't attend include (i) some farmers have side-businesses and have to tend to them (ii) they may have jobs in other locations (iii) some farmers were discouraged as they were recently cheated of 10kgs at the wet-mill. If a farmer does not attend twice in a row, the farmer is reported to the focal farmer who follows up with the farmer for the reasons behind his/her absence. Overall, Deo is very satisfied with the training. All the farmers know each other well but they don't really meet beyond the training. The best training was on nutrition (using NPK) and the different types of composting – he had never seen such techniques.

Farmer Trainers

Before 2 men from TNS would come, one to measure the altitude, latitude, and other geographical characteristics, whereas the other TNS person would measure BP adoption. Sometimes the NAEB official would make a check on the coffee too – atleast twice a month but it has been 5 months since he last came. Usually, Deo is not aware when the farmer trainer is coming to check his coffee trees but knows he comes regularly which motivates him to check on his fellow farmers' fields "*when the superior is coming, it reminds me to check on other farmers*"[1].

Farmer Profiles: Silvette & Fidel

Personal

Silvette is a 43 year old farmer with 3 children who are all in secondary school. Her husband is not usually around so she handles all the farming responsibilities and attends all the TNS trainings.
Fidel is a 40 year old married farmer with 5 children. His wife helps him with the farming, especially with mulching; he also has 2 hired workers that help him with the weeding. His wife also alternates attending the training as he has lots of meetings so cannot attend trainings all the time.

They are both members of the cooperative and most of the famers they know are also members. The tend to pool the crop together and have regular meetings at the cooperative level to discuss the training (cooperative meetings are three times/year).

Informed of training?

For Fidel and his fellow farmers, the farmer trainer informs the umudugudu head who goes around the neighborhood "shouting" the training date. The farmer trainer or data collector visit atleast once a month to check on his farms and make sure that the BP training is implemented. Sometimes he is informed that the farmer trainer as the focal farmer will be alerted. The focal farmer then informs him that the FT will be coming to visit his farm; sometimes the FT will drop by unexpected and tell him how to improve his techniques.

Silvette lives far from the focal farmer and other farmers so she usually gets a personal phone call from the focal farmer. She usually encounters him three times a week on her way to the farm. As a result the focal farmer makes regular visits to her farm to check on the BPs. The farmer trainer also visits once a month to check the trees and for diseases.

Knowledge

Before TNS, agronomists from the district would provide training. They had some knowledge of coffee farming from their school days and also from their parents and families but they have not had training from any other source.

TechnoServe training

Members tend to miss training due to sick family members, other meetings, and the distances involved in traveling (e.g. takes 20mins to get to their farms and 40mins to get to focal farmer for training!). Before the TNS training, they really did not have much knowledge about these techniques and it has resulted in an increase of yields. The best aspects of the training include mulching, composting, IPM, and how to set up a cooperative. Recommendations for improvement include the provisions of free NPK and pesticides, more regular training session and more farmer trainers to provide training and checking in on them.

Farmer Trainers

At the training, the FT encourages participants to attend the training. Silvette admits that having the farmer trainer around acts as a reminder for her to implement the BPs she's learnt.

Annex B: Technical Annexes

Testing the effect of time

Estimating t1 (linear regression)

	Dependent variable : Yield (Unit: values in kgs/tree)				
Variable	**1**	**2**	**3**	**4**	**5**
t1 estimate (explanatory)	-0.007	-0.001	0.016	-0.050	-0.035
Number of trees owned by farmer		0.000***	0.000***	0.000***	0.000***
Dummy for female		0.170*	0.169*	0.169*	0.164*
Dummy for cooperative member		-0.026	0.030	0.030	0.024
Average altitude in cooperative			0.000	0.000	0.000
Total rainfall in cooperative (2010)			0.001***	0.001***	0.001*
Share of productive trees in sector				-0.709*	-0.645
Poverty level in district					-0.300
Constant	1.677***	1.739***	0.571	0.982	1.302
Number of observations	347	347	347	347	347

Asterisks denote the following significance level *p<0.01 ** p<0.05 ***p<0.01
Green, light yellow and light orange respectively also denote *p<0.01 ** p<0.05 ***p<0.01

Estimating t2 (linear regression)

	Dependent variable : Yield (Unit: values in kgs/tree)				
Variable	**1**	**2**	**3**	**4**	**5**
t2 estimate (explanatory)	0.035	0.000	-0.027	-0.144	-0.088
Number of trees owned by farmer		-0.001**	-0.001**	-0.001***	-0.001***
Dummy for female		0.084	0.092	0.090	0.087
Dummy for cooperative member		0.028	0.042	0.051	0.042
Average altitude in cooperative			-0.001	0.000	-0.001**
Total rainfall in cooperative (2011)			0.000	-0.002	-0.001
Share of productive trees in sector				0.864	0.772
Poverty level in district					-0.921
Constant	1.670***	1.772***	3.211**	4.007**	5.008***
Number of observations	353	353	353	353	353

Asterisks denote the following significance level *p<0.01 ** p<0.05 ***p<0.01
Green, light yellow and light orange respectively also denote *p<0.01 ** p<0.05 ***p<0.01

Estimating the impact of the program

Estimating impact of program on Cohort 2010 after two years of training (linear regression)

Treatment Group = Cohort 2010 (Follow-up data after 2 years of training)
Control Group=Cohort 2012 (Baseline)

	Dependent variable : Yield (Unit: values in kgs/tree)				
Variable	**1**	**2**	**3**	**4**	**5**
Treatment (explanatory)	1.316***	1.337***	1.242***	1.239***	1.234***
Number of trees owned by farmer		-0.001***	-0.001***	-0.001***	-0.001***
Dummy for female		-0.214	-0.196	-0.197	-0.198
Dummy for cooperative member		0.085	0.092	0.094	0.092
Average altitude in cooperative			-0.002**	-0.002**	-0.002**
Total rainfall in cooperative (2010)			0.000	0.000	0.000
Share of productive trees in sector				0.089	0.095
Poverty level in district					-0.059
Constant	3.021***	3.369***	6.561***	6.516***	6.574***
Observations	326	326	326	326	326

Asterisks denote the following significance level *p<0.01 ** p<0.05 ***p<0.01
Green, light yellow and light orange respectively also denote *p<0.01 ** p<0.05 ***p<0.01

Estimating increase in yields in Cohort 2010 after one year of training (linear regression)

Treatment Group = Cohort 2010 (Follow-up data after 1 years of training)
Comparison Group=Cohort 2011 (Baseline)

	Dependent variable : Yield (Unit: values in kgs/tree)				
Variable	**1**	**2**	**3**	**4**	**5**
Treatment (explanatory)	1.009***	0.961***	0.967***	0.990***	0.958***
Number of trees owned by farmer		-0.001	-0.001	-0.001	-0.001***
Dummy for female		0.007	0.004	0.004	-0.005
Dummy for cooperative member		-0.047	-0.047	-0.047	-0.059
Average altitude in cooperative			0.000	0.000	0.000
Total rainfall in cooperative (2010)			0.000	0.000	0.000
Share of productive trees in sector				-0.237	-0.100
Poverty level in district					-0.640
Constant	2.679***	2.865***	2.600***	2.736***	3.419***
Observations	347	347	347	347	347

Asterisks denote the following significance level *p<0.01 ** p<0.05 ***p<0.01
Green, light yellow and light orange respectively also denote *p<0.01 ** p<0.05 ***p<0.01

Estimating increase in yields in Cohort 2011 after one year of training (linear regression)

Treatment Group = Cohort 2011 (Follow-up data after 1 years of training)
Comparison Group=Cohort 2012 (Baseline)

	Dependent variable : Yield (Unit: values in kgs/tree)				
Variable	**1**	**2**	**3**	**4**	**5**
Treatment (explanatory)	1.193***	1.165***	1.164***	1.395***	1.343***
Number of trees owned by farmer		-0.001***	-0.001***	-0.001***	-0.001***
Dummy for female		-0.119	-0.111	-0.115	-0.117
Dummy for cooperative member		0.351***	0.411***	0.427***	0.419***
Average altitude in cooperative			-0.001	-0.001	-0.001
Total rainfall in cooperative (2010)			0.001	-0.002	-0.002
Share of productive trees in sector				1.708***	1.622***
Poverty level in district					-0.870
Constant	2.899***	3.030***	4.203	5.776**	6.723**
Observations	353	353	353	353	353

Asterisks denote the following significance level *p<0.01 ** p<0.05 ***p<0.01
Green, light yellow and light orange respectively also denote *p<0.01 ** p<0.05 ***p<0.01

Highlighting the link between attendance and Best Practice Adoption

Link between attendance and number of best practices adopted (logit regression)

Data set = Cohort 2010, round 2 best practice data collected in the second half of 2011

	Dependent variable: X BPs adopted					
Variable	**6 BPs**	**7 BPs**	**8 BPs**	**9 BPs**	**10 BPs**	**11 Bps**
Attendance rate (explanatory)	2.496***	2.558***	2.094***	1.848***	1.613***	2.525**
Dummy for female	0.074	-0.332	0.064	0.033	-0.055	0.207
Dummy for cooperative member	-0.164	0.071	0.341*	0.553	0.123	-0.223
Dummy for focal farmer	-0.130	0.368	1.596**	0.533*	0.552*	1.186*
Focal farmer group size	-0.048	-0.049*	-0.009	-0.001	-0.006	-0.027
Cooperative dummies	Yes	Yes	Yes	Yes	Yes	Yes
Constant	3.293*	1.538	-0.286	-1.348**	-1.817*	-3.583**
Number of observations	562	741	832	832	832	793

Asterisks denote the following significance level *p<0.01 ** p<0.05 ***p<0.01
Green, light yellow and light orange respectively also denote *p<0.01 ** p<0.05 ***p<0.01

Link between fertilizer usage and best practices adopted (logit regression)

Data set = Cohort 2010, round 2 best practice data collected in the second half of 2011

	Dependent variable: dummy for type of fertilizer			
Variable	**Composting**	**NPK**	**Lime**	**Zinc/borium**
Attendance rate (explanatory)	0.841	1.114	3.828***	2.863***
Dummy for female	0.242	0.049	-0.692*	-0.678
Dummy for cooperative member	0.731**	0.559	-0.573	-0.271
Dummy for focal farmer	0.765	1.509***	1.395	2.569
Focal farmer group size	0.063	0.004	-0.106	0.006***
Cooperative dummies	Yes	Yes	Yes	Yes
Constant	-3.089*	-3.632***	1.243	-5.879***
Number of observations	826	832	403	647

Asterisks denote the following significance level *p<0.01 ** p<0.05 ***p<0.01
Green, light yellow and light orange respectively also denote *p<0.01 ** p<0.05 ***p<0.01

Estimating the impact of monitoring

Impact of belonging to yield sample on attendance and best practice adoption (linear regression)

	Dependent variable:	
Variable / number of BPs adopted	**Attendance**	**Total BP adoption**
Dummy for yield sample	0.073	0.580
Dummy for female	0.024	0.014
Dummy for cooperative member	0.035	0.182
Dummy for focal farmer	0.082	0.568
Focal farmer group size	-0.001	-0.007
Cooperative dummies	Yes	Yes
Constant	0.869	8.864
Number of observations	1029	1029