IMPACT STUDY

Beyond Perception

The Critical Role of Attribution in Assessing Mobile Agronomic Services for Kenyan Macadamia Process

Juan Taborda, Jorge Bouchot, and Miranda Hansen
245 Main Street, 2nd Floor,
Cambridge, MA 02142
United States
jtaborda@rootcapital.org

Root Capital

CONTENTS

ACKNOWLEDGEMENTS

This study was possible thanks to the willingness of our clients and their farmers to participate in this service and evaluation. Their leadership and their employees played a vital role in supporting the service implementation and data collection efforts.

This evaluation draws upon the work conducted by staff and collaborators at Root Capital who contributed to the implementation of services and the execution of the present evaluation. They also identified key partners to fund and support the completion of this report. The list of collaborators includes all those who worked on implementing the DEAS service in Kenya, particularly our Africa Advisory team, including Rose Ochieng, Washington Ndwiga, Alain Bunani, and Beth Wangui, and our Global Advisory colleague, Lia Hulit. Colleagues on the Communications, External Affairs, Client Services, and Strategy teams in Africa and at the global level were also instrumental in the completion and review of this evaluation.

We would also like to acknowledge the significant contributions of former colleagues who were instrumental in the early stages of this project, including Juana Muñoz and Jimmy Vulembera.

Finally, we would like to extend our gratitude to the Swiss Re Foundation, whose partnership and generous support for our work made this project possible.

EXECUTIVE SUMMARY

Study Motivation and Approach

This report evaluates the impact of Digital Extension Advisory Services (DEAS) on macadamia farmers in Kenya. While our findings suggest that DEAS delivered value to its recipients, our causal assessment indicates that the program did not achieve its intended outcome of increasing farmers' knowledge of good agricultural practices.
The primary objective of this study was to determine if the DEAS model, which has demonstrated some success in increasing the knowledge and adoption of good agricultural practices among coffee farmers in Rwanda, could be effectively replicated in a different agricultural sector and geographic context. Building upon the learnings of Root Capital's previous DEAS interventions in Rwanda, this research seeks to assess the transferability of the service and its potential to generate similar benefits for macadamia farmers in Kenya. It also intends to provide better insights on the validity of the model by using a randomized controlled trial to unequivocally assess the impact of the service.

As work on the current evaluation progressed, it became clear that the learning generated went beyond simply analyzing the impact of the DEAS service to allowing us to engage in a more meaningful conversation about how to better assess the impact of programs and services in the field. The results on this iteration of DEAS are meaningful as they will inform programmatic decisions around this and other similar services. However, the key conversation we would like to advance is related to how organizations like Root Capital should be careful not to misinterpret respondent perceptions about their services with the actual effects attributed to their programs.

An initial assessment of farmers' perceptions of the service suggests a positive program outcome. The Net Promoter Score (NPS®) derived from survey responses was favorable (n=559; score of 6.32), with more farmers categorized as "promoters" than "detractors" of the service. This means respondents of the survey would recommend the service to family or friends. However, a more detailed analysis of learning outcomes among participants with both baseline and endline data, as well as comparing randomized treated farmers versus control groups, reveals a different story. Despite the positive NPS®, the Randomized Controlled Trial (RCT) and the Difference-in-Differences analysis demonstrates that there is no significant effect of the service on learning outcomes.

This nuanced finding underscores the importance of rigorously isolating the impact of a service like DEAS. While the initial positive perception on behalf of beneficiaries may be encouraging, it is crucial to conduct a thorough evaluation to determine the actual effectiveness of the program in achieving its intended goals.

Main Findings

In October 2023, after an implementation with coffee cooperatives in Rwanda, Root Capital expanded DEAS to agricultural processors in the macadamia nut value chain in Kenya with farmers in the supply chain of three processors. The provision of this service was designed to last no more than one year with modules addressing topics related to the adoption of good agricultural practices (GAPs) in macadamia nut production.

The intervention included a baseline SMS-based survey to assess knowledge about GAPs. This was followed by 37 SMS reminders encouraging farmers to adopt specific agricultural practices that aligned with the stage of their production cycle. At the conclusion of the service, producers received a follow-up survey, this time in Interactive Voice Responses (IVR) format to increase response rates. This survey included the same questions as the baseline survey, along with two additional questions: one assessed the satisfaction of the service, and the other asked farmers if they remembered receiving messages with agronomic guidance.

We faced three main challenges in delivering the service and conducting the surveys: attrition, low response rates, and limited take-up (the rate at which people engage with a service). A total of 6,042 farmers were registered to receive the service. During the program implementation, 1,301 farmers (22%) requested their network providers to unsubscribe their numbers from mass SMS campaigns, suggesting they may not have received some or all of the messages. For the surveys, 8.3% of the farmers responded to the baseline survey, and response rates increased to 19% with the endline IVR survey. Finally, only 335 farmers recalled receiving the training materials, indicating a 5.7% take-up rate among those registered for the service.

To rigorously estimate the effect of receiving the SMS messages, farmers were randomly assigned to three groups: all receiving core messages and two groups receiving only partial content. Given that the survey covered all training topics, it was expected that participants who received all the content would have higher knowledge scores about GAPs than those who received only partial content.

Our regional team reported that all macadamia processors and most of their affiliated farmers were satisfied with DEAS. Consistent with this, most farmers who responded to the final survey indicated they would recommend the service to a friend or relative, resulting in an NPS®-like score of 6.32. However, when analyzing changes in knowledge scores between baseline and endline, or significant treatment effects between groups, we found no evidence of learning outcomes. These null results indicate that the training did not achieve the expected learning effects observed in other industries and countries. Our randomization was effective, so we should have observed a difference between fully treated and partially treated groups if significant learning effects were present. We also have no reason to believe that we were underpowered to detect statistically significant effects based on previous results in other countries or industries.

INTRODUCTION

Study Objective

This report evaluates the impact of the digital extension advisory services (DEAS) on macadamia farmers in Kenya. The primary objective of this study was to determine if the DEAS model, which has demonstrated some success in increasing the knowledge and adoption of good agricultural practices of farmers affiliated to coffee cooperatives in Rwanda[1], could be effectively replicated in a different agricultural sector and geographic context. It also intended to provide better insights on the *true effects* of the service by using a randomized controlled trial to isolate and unequivocally assess its impact.

INTERVENTION – DEAS IN MACADAMIA ENTERPRISES IN KENYA

Providing extension services to farmers via Short Message Service (SMS), Interactive Voice Responses (IVR) and voice calls have the potential to increase farmers' knowledge and use of good agricultural practices (GAPs) in a cost-effective manner. Since 2019, Root Capital has been providing DEAS to farm households associated with some of our clients in Africa.

Root Capital initiated DEAS in 2019 for Rwandan coffee farmers associated with two of our clients. An initial evaluation confirmed program utilization, knowledge gain, and behavioral changes among farmers (Naeve et al. 2022). This included increased awareness of GAPs and improved application of nearly all targeted practices. However, a second evaluation found no statistically significant impact on knowledge or GAP application from extending DEAS for an additional year of service (Taborda and Busto 2023). Informed by these findings, Root Capital modified the implementation of the DEAS service, restricting the service to only one year of duration, given the second evaluation suggested the knowledge gained by DEAS plateaued after a year of implementation.

In October 2023, Root Capital extended DEAS to farmer members of three agricultural processors in the macadamia nut value chain in Kenya. The provision of this service was designed to last no more than one year, with seven modules addressing topics related to the adoption of GAPs in macadamia nut production. It is also important to note that due to the role

[1] See Naeve et al. (2022) and Taborda & Busto (2023).

Root Capital's clients have in the production and processing of macadamia nuts, there is a group of farmers who received the intervention who are likely linked to more than one client.

The sociodemographic descriptive statistics of the farmers participating in the program are presented in Table 1. Approximately one-third of participants are women, and the gender and geographic distribution of participants remained largely unchanged between the two data collection periods (baseline and endline).

Table 1: Farmers' Sociodemographic Descriptive Statistics

Variable	Baseline (n=6,042)		Endline (n=5,927)	
	n	%	n	%
A. Gender				
Female	1,947	32.23	1,914	32.31
Male	4,094	67.77	4,010	67.69
B. Geographic distribution: Counties				
Baringo County	62	1.03	61	1.03
Chuka County	165	2.73	165	2.79
Embu County	3,942	65.24	3,829	64.66
Kiambu County	120	1.99	120	2.03
Kirinyaga County	91	1.51	88	1.49
Machakos County	145	2.40	145	2.45
Meru County	637	10.54	636	10.74
Muranga County	502	8.31	501	8.46
Nyeri County	378	6.26	377	6.37

The intervention consisted of a baseline SMS-based survey to assess knowledge about GAPs, followed by 37 SMS reminding farmers to adopt specific agricultural practices in line with the corresponding period of the nut production cycle.

In collaboration with our regional experts in agronomic extension for macadamia and our global agronomic and climate resilience advisory experts, we identified a core set of practices that all farmers should be provided alongside two sets of additional practices that could potentially improve their productivity, but were not considered essential. This allowed the evaluation design to restrict the content provided to some farmers without presenting an ethical concern of limiting access to fundamental information. Farmers were assigned to one of three groups (0, 1, and 2, see Table 2 below) through a stratified random assignment, with the commitment by the team to share all the content information if the evaluation validated a significant impact of the service on farmers' knowledge about these additional practices. Group 0 received comprehensive advisory guidance on all nine agricultural practice topics. Groups 1 and 2 received a reduced scope of information, focusing on only six of the nine advisory topics. Group 1 received 25 messages, and Group 2 received 28 messages.

Table 2: Evaluation Design – Allocation of advisory topics by groups

Topics	Group 0	Group 1	Group 2
1. Soil fertility management	✓	✓	✓
2. Harvesting only from ground	✓	✓	✓
3. Dehusking after harvesting	✓	✓	✓
4. Altitude specific variety	✓	✓	
5. Soil erosion control	✓	✓	
6. Intercropping different species	✓	✓	
7. Pruning	✓		✓
8. Integrated pest and disease	✓		✓
9. Mulching for soil and water management	✓		✓

At the conclusion of the service in August 2024, the same baseline survey was conducted using an IVR methodology, which involved completing the survey through phone calls rather than over text messages. The decision to shift towards IVR was motivated to increase the response rate of the survey, based on the experience the Africa Advisory Team had in recent iterations of DEAS in Rwanda. As detailed in the data section, this new mechanism significantly improved survey response rates in the endline compared to the baseline survey, showing an increase by 10 percentage points for all questions and by 7 percentage points for at least one survey question.

SMS or IVR interventions

The widespread adoption of mobile phones has significantly contributed to the increasing prevalence of SMS interventions within the international development sector. One of the primary reasons for this trend is the substantial cost reduction associated with SMS interventions compared to more intensive behavioral change programs.

In addition to Root Capital's first-hand experience in assessing the impact of SMS interventions in agriculture, academic literature has identified positive impacts of this type of behavioral intervention in agriculture (Aker, Ghosh, and Burrell 2016). (Aker et al. 2021) reviewed 61 studies using SMS interventions across five policy relevant areas: agriculture, education, electoral outcomes, financial inclusion, and public health. Their analysis draws attention to three findings: i) the impact of SMS interventions is mixed, dependent upon the intervention, country, sector, and sample size of the study; ii) imperfect compliance is a systematic issue in SMS interventions, ranging considerably across studies, from 1.3% to 98% of take-up rates being reported; iii) low compliance could potentially explain null results in studies; and has important implications for the interpretation of intervention impacts.

Of the studies reviewed in this meta-analysis, nearly 70% of interventions show positive effects, but only half show significant effects for most treatment arms. For the studies reviewed in agriculture, all show significant effects, but most studies have additional interventions that are coupled with the SMS or IVR reminder.

Learning from the challenges identified in the Aker et. al.'s (2021) meta-study, this report will not only focus on the issue of compliance by providing different measurements of take-up (both for the SMS baseline survey and the IVR endline survey) but will also estimate treatment effects using different econometric techniques to assess the generalizability of the results. For more on these topics, see section two of the Technical Appendix.

Evolving Dimensions of the Evaluation

When designing the evaluation, the research team initially focused on navigating imperfect compliance, when not all participants follow the assigned treatment protocol, with SMS interventions. To address this, the evaluation design included not only a baseline survey (which is typically unnecessary for a RCT design), but also the estimation of both intention-to-treat effects and local average treatment effects. This allowed us to assess the impact of both being assigned to the intervention and actually receiving it. However, as we analyzed the data, some valuable insights emerged that we could offer to the overall discussion about the usage of perception-based measures and the use of econometric methods to make the claim of attributable impacts.

After a thorough analysis of the endline data, discrepancies emerged. The 559 responses provided in the endline survey indicated a high level of satisfaction among participants. Overall, NPS® for the service was positive, with a score of 6.32. This suggests that a larger number of respondents would recommend the service to family or friends compared to those who would not. However, the impact estimated by all econometric methods utilized did not show significant learning outcomes for improvement of knowledge for GAPs targeted by the intervention.

This dichotomy between the perception of farmers who responded to the survey and would recommend the service to their friends or relatives, with their own performance on the assessment about knowledge of agricultural practices, highlighted the importance of rigorously measuring the impacts of our service beyond perceived improvements.

Over the last few years, we have seen in our field an increase in the use of perception-based measures to claim improvements in outcomes and quantify impact. While these tools and methods provide valuable information and help us better understand how services are received in the field or how firms are perceived compared to peers in similar markets, it is crucial to assess not only the perceptions of service recipients but also have objective measures of the intended outcomes of the programs or services. Clearly, DEAS provided value for survey respondents, but after rigorous assessment, it appears that the program did not achieve its intended outcomes in increasing farmers' knowledge on good agricultural practices.

METHODOLOGY

The evaluation employed a RCT to measure the effect of DEAS on knowledge around agricultural practices. The study aimed to estimate both the intention to treat effect (impact on all participants assigned to receive the service or ITT) and the local average treatment effect (LATE), which focuses on those who actually received the service, using the take-up rate as an instrumental variable. This design intended to address the issue of imperfect compliance, a common concern in studies of SMS or IVR interventions in development economics.

The evaluation also collected baseline and endline data to assess whether there were observable differences among participants using a Difference-in-Differences approach, a statistical method that compares changes over time between a treatment group and a control group, to ensure we had robust alternatives to estimate the effects of the service. Despite having a sample size that was considered adequate for detecting significant effects, no statistically significant differences were observed.

Identification Assumption – Successful Randomization

The fundamental assumption of an RCTs is that randomization effectively balances treatment groups on all relevant baseline characteristics. This ensures that any observed differences between groups can be attributed only to the intervention. Tables 3 and 4 demonstrate the balance achieved in our study and present baseline characteristics, such as gender and county of residence, that should not be affected by the intervention. As shown, there are no significant differences between the groups for any of these characteristics, indicating that the randomization process was successful.

To ensure the randomization was effective, it was essential to verify that there were no issues with non-compliance (when participants do not follow the assigned treatment) or differential attrition (when participants drop out of the study at different rates across groups). These issues could lead to unbalanced groups and undermine the validity of the randomization. Additionally, given that we were using an SMS-IVR intervention, it was important to assess these factors to accurately estimate the treatment effects. The following subsection details the response rates for all intervention groups and the differences in response rates between the two rounds and methods of data collection.

Table 3: Balance Table of Baseline Characteristics

Variable	Group 0		Group 1		Group 2		Means Differences	
	n_0	Mean (0)	n_1	Mean (1)	n_2	Mean (2)	=(0)-(1)	=(0)-(2)
A. Gender								
Female (%)	2,018	32.11	2,011	32.27	2,008	32.27	-0.16	-0.16
B. Geographic distribution: Counties								
Baringo County (%)	2,018	0.84	2,011	0.99	2,008	1.20	-0.15	-0.35
Chuka County (%)	2,018	3.17	2,011	2.64	2,008	2.39	0.54	0.78
Embu County (%)	2,018	65.01	2,011	65.39	2,008	65.49	-0.38	-0.47
Kiambu County (%)	2,018	1.83	2,011	2.04	2,008	2.09	-0.21	-0.26
Kirinyaga County (%)	2,018	1.19	2,011	1.74	2,008	1.54	-0.55	-0.35
Machakos County (%)	2,018	2.28	2,011	2.19	2,008	2.74	0.09	-0.46
Meru County (%)	2,018	10.90	2,011	10.09	2,008	10.56	0.81	0.34
Muranga County (%)	2,018	8.37	2,011	8.40	2,008	8.17	-0.03	0.21
Nyeri County (%)	2,018	6.39	2,011	6.51	2,008	5.83	-0.12	0.57

Table 4: Balance Table of Endline Characteristics

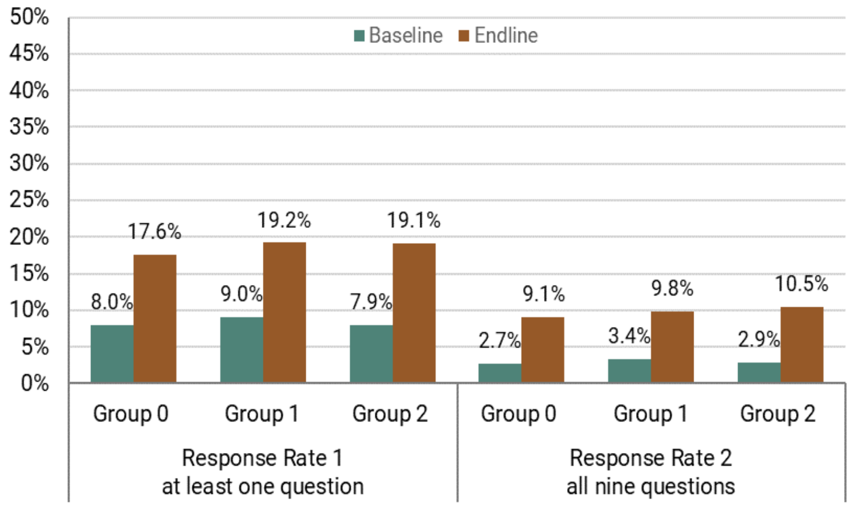
Variable	Group 0		Group 1		Group 2		Means Differences	
	Obs	Mean (0)	Obs	Mean (1)	Obs	Mean (2)	=(0)-(1)	=(0)-(2)
A. Gender								
Female (%)	1,970	32.54	1,973	32.44	1,980	31.97	0.10	0.57
B. Geographic distribution: Counties								
Baringo County	1,968	1.02	1,973	0.96	1,980	1.11	0.05	-0.09
Chuka County	1,968	2.74	1,973	2.79	1,980	2.83	-0.04	-0.08
Embu County	1,968	64.33	1,973	64.98	1,980	64.65	-0.65	-0.32
Kiambu County	1,968	2.13	1,973	1.98	1,980	1.97	0.16	0.16
Kirinyaga County	1,968	1.52	1,973	1.47	1,980	1.46	0.05	0.06
Machakos County	1,968	2.49	1,973	2.43	1,980	2.42	0.06	0.07
Meru County	1,968	10.72	1,973	10.75	1,980	10.76	-0.02	-0.04
Muranga County	1,968	8.59	1,973	8.36	1,980	8.43	0.22	0.15
Nyeri County	1,968	6.45	1,973	6.28	1,980	6.36	0.17	0.09

Response rates and Knowledge Scores for GAPs

Overall, 6,042 farmers were registered to receive DEAS. Of these, 502 (8.3%) responded to at least one question from the SMS baseline survey, and 180 (3%) responded to all nine questions. After nine months of service implementation, an endline survey was delivered to 5,927 farmers. The delivery report indicates that 4,630 contacts received all the training messages they were assigned. Of these, 1,124 farmers (19%) responded to at least one question of the IVR survey, and 591 (10%) responded to all nine questions.

A key finding from this evaluation, as evidenced in Figure 1, was that IVR surveys significantly outperformed SMS-only surveys in terms of response rates. We observed an increase of over 10 percentage points in response rates for at least one question, and 7 percentage points in the rate of completing all survey questions when using IVR compared to SMS.

Figure 1: Survey Response Rates per Round of Data Collection

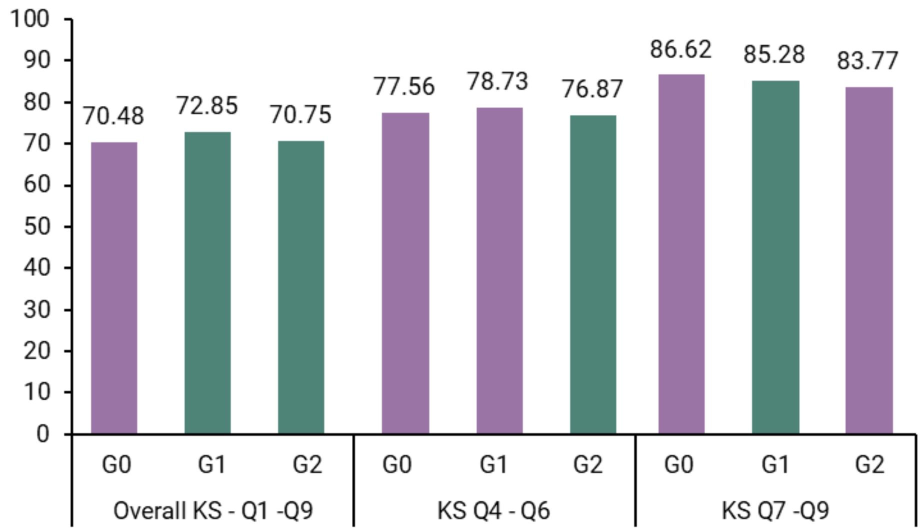


To quantify the learning outcomes of the program, we calculated a Knowledge Score (KS) for good agricultural practices, defined as the percentage of correct answers with respect to the number of questions answered. In order to explore the differences in learning outcomes by allocation of advisory topics (see Table 2 above), we created three different scores:

- An overall knowledge score for the nine questions of the surveys.
- A knowledge score only for questions 4 to 6 in the survey (Group 2 did not receive these advisory topics before the endline survey).
- A knowledge score for questions 7 to 9 in the survey (Group 1 did not receive these advisory topics before the endline survey).

Figure 2 illustrates that there is essentially no difference in knowledge scores among the three groups. Given that farmers in Group 0 received all the agronomic advisory topics, we would expect them to have a higher score compared to the other two groups. However, the graph shows that farmers in Group 0 had a slightly lower overall score (70.48) than those in Group 1 (72.85) and Group 2 (70.75), but it is important to emphasize these differences are not statistically significant. The other two knowledge scores (Q4-Q6 and Q7-Q9) also show only marginal differences, suggesting that was no statistically significant difference between receiving or not specific agronomic advisory topics. This observation is further supported by a causal estimation of the impact of DEAS on learning outcomes, as detailed in the Findings section below

Figure 2: Average Knowledge Scores (KS) for GAPs by group and by questions included for its calculation



Note. Green bars represent groups that did not receive all topics, so a lower KS was expected. Purple bars represent the group(s) for which a higher KS was anticipated.

NPS® score and perceptions of service success

One of the most common metrics used to assess client satisfaction with a service is the Net Promoter Score (NPS®)[2]. This tool, designed by Bain & Company Inc., has proven to be extremely reliable and effective in gauging client perception of a company's service provision. The traditional design uses an 11-point scale (ranging from 0 to 10), which provides sufficient granularity for respondents to be assigned to one of two categories: promoters or detractors.

[2] Information on the NPS® and its estimation in its traditional framework comes from Bain & Company Inc (Bain & Company Inc. 2017).

The NPS® is calculated from answers to the question "*How likely are you to recommend [Company/Service] to a friend or colleague?"*. The score is the percentage of customers who are "promoters" of the company or service (answers of 9 or 10) minus the percentage of "detractors" (those who answered between 0 to 6).

While the original NPS® relies on an 11-point scale since its formulation in 2003, researchers have adjusted the scale to a 5-point version to reduce input errors, especially when using SMS or IVR based questions that require keypad input. Analyses suggest that changing from an 11-point to a 5-point scale makes the tool slightly less precise, with average absolute errors of about 4 percentage points, but still relatively reliable for estimating customer satisfaction (Sauro 2017). Field implementation revealed that with the target population participating in the study, typing errors on the scale led to significant concerns regarding data reliability. Following the recommendation of the service implementer, we opted for the 5-point scale over the traditional 11-point alternative.

The assessment of farmers' perceptions of the service using this NPS®-like measure suggests a positive program outcome, or more carefully worded, a good perception of farmers' experience with the DEAS service. The NPS®-like score derived from survey responses was favorable (n=559; score of 6.32), with more farmers categorized as promoters than detractors of the service, implying that most farmers would recommend the service to friends or colleagues.

Pre-Post score comparison

Another alternative that has been used in the field to assess the success of a program implementation is a before-and-after performance evaluation. The idea behind this methodology is comparing outcomes of interest in two periods of time: prior the implementation of the intervention and once the intervention was finalized. However, the main challenge of this methodology is that it does not allow to isolate the causal effect of the intervention in a way in which it is possible to identify the attributable impact of a program like DEAS.

When applying this methodology to the 39 participants with both baseline and endline survey data, we found that 11 respondents (28.2%) demonstrated improved knowledge in at least one of the nine questions. However, as noted previously, a limitation of this approach is its inability to definitively attribute these improvements solely to the DEAS intervention.

Take-up and Attrition

Take-up and attrition are two major concerns that highlight significant issues in program implementation for behavioral phone or SMS interventions. When measuring the take-up rate of the service, one could rely exclusively on reports of successful delivery of SMS to farmers. Overall, 5,927 farmers were registered to receive the service in September 2023. Over the nine months of service implementation 1,301 farmers requested their telecommunications providers to unsubscribe their numbers from mass SMS campaigns, suggesting they may have not received some or all the messages. The delivery report indicates that 4,630 contacts received all training messages they were assigned to receive. Therefore, relying exclusively on the successful delivery of the service will lead us to estimate a take-up rate of 78.1%.

However, given the existing literature on SMS interventions, we decided to ask respondents whether they remembered receiving training materials through SMS messages at the conclusion of the program. Of the 546 respondents who answered this question, only 61.4% remembered receiving the training materials. Given the rate of response of the survey compared to the potential population that received the service, we estimate our take-up rate of the service was 5.7% of the entire population registered to receive the service, or 7.2% of the sample to whom the SMS were successfully delivered during the service implementation.

To account for non-compliance in the assignment to the treatment, as evidenced by all our take-up estimates, we also used an instrumental variable method to estimate the LATE and assess the effectiveness of the intervention while controlling for partial compliance for the farmers who successfully received the SMS. For more on this estimation, please review the Technical Appendix.

Regarding differential attrition, we have no concerns about a systematic difference in attrition between our treatment groups, suggesting that the randomization was successful. However, despite not being systematically different, we are deeply concerned about the alarmingly high attrition rates when comparing participants in the baseline survey to respondents of the endline survey. Only 39 of the 502 respondents that answered at least one question in the baseline also responded to at least one question in the endline, corresponding to an attrition rate of 92%. Similarly, only 13 of the 180 respondents who answered all nine questions in baseline also answered the whole survey in the endline, corresponding to an attrition rate of 92.8%.

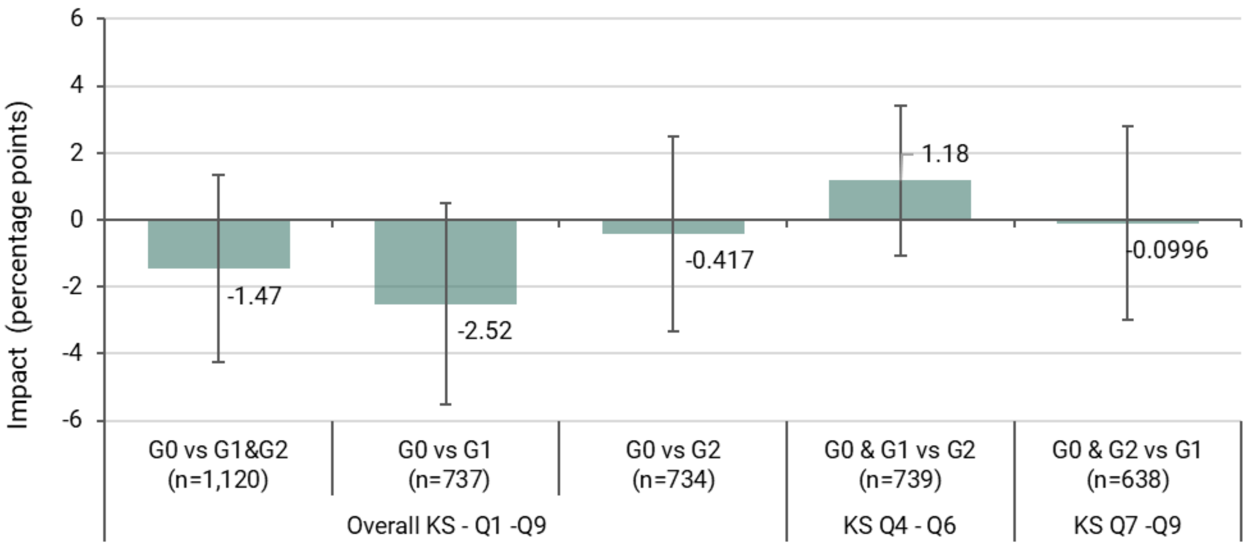
FINDINGS

The NPS® data shows positive farmer satisfaction with DEAS. However, despite using various estimation methods, we found no statistically significant effect of DEAS on GAP knowledge among macadamia farmers in Kenya. This means that, on average, farmers who received more training material through SMS messages did not improve their knowledge of GAPs compared to those who received limited training.

Figure 3 and Table 5 present the estimations of the ITT method[3], which was the simplest and least restrictive method to determine a statistically significant effect of our intervention.

The results obtained using this estimation strategy are consistent with those from all other alternative estimation methods, including the instrumental variable method to account for imperfect compliance and the Difference-in-Differences method tested as a robustness check. Additionally, when comparing the results to previous studies and findings in the literature, we have no reason to believe that we are underpowered to detect statistically significant effects.

Figure 3: Regression Results – Impact Knowledge Score (KS), different comparison groups (percentage points)



Note. The bars indicate the estimation of DEAS' impact using different comparison groups. The whiskers are their respective confidence intervals at the 95%, showing that in all cases the impact is not statistically different from zero.

[3] For more on the econometric model about this model please see Technical Appendix.

Table 5: Regression Results – RCT / ITT

Variables	Overall Knowledge Score (KS) - Q1 -Q9			KS Q4 - Q6	KS Q7 -Q9
	(1) G0 vs G1&G2	(2) G0 vs G1	(3) G0 vs G2	(4) G0 & G1 vs G2	(5) G0 & G2 vs G1
a) RCT - ITT					
DEAS Intervention	**-0.0147**	**-0.0252**	**-0.00417**	**0.0118**	**-0.000996**
	(0.0143)	**(0.0153)**	**(0.0148)**	**(0.0224)**	**(0.0148)**
Constant	0.718***	0.729***	0.707***	0.769***	0.853***
	(0.0119)	(0.0112)	(0.0145)	(0.0142)	(0.0122)
Observations	1,120	737	734	739	638

Asterisks denote statistical significance at the confidence interval of 99% (***), 95% (**) and 90% (*), respectively.

Alternative Mechanisms

Given the evidence presented, we are confident in stating that for this iteration of DEAS, there might be some structural issues present that resulted in the service not showing similar gains in knowledge about agricultural practices compared to previous evaluations we or others have conducted on these types of intervention in the agricultural sector. While we did not fully explore some of these alternatives through rigorous methods, we do believe there are two characteristics of the way the service was implemented and the macadamia industry in Kenya that may explain the null results observed.

First, it is important to note that this is the first evaluation by Root Capital which isolates the effect of DEAS from other Root Capital programs (due to the RCT design). This is also the first time we have evaluated the service in an industry where our clients are more likely to act as aggregators (macadamia nut processors) than as collaborators (coffee cooperatives in Rwanda) with close relationships with their supply chains. In previous evaluations, either due to the industry structure or concurrent services by the cooperatives, the SMS intervention was not the only potential factor affecting improved knowledge of GAPs. Given the cooperative structure of other clients who participated in previous DEAS evaluations (all coffee cooperatives in Rwanda), one can imagine how the service, combined with more active technical assistance from the cooperatives, led to the messages being reminders or reinforcements of knowledge conveyed through other mechanisms, reinforcing both knowledge and adoption of GAPs.

Second, it is important to note that country-specific issues might suggest farmers are overwhelmed by SMS or IVR messages. One element suggests an oversaturation of our sample with mobile-based interventions, while the other hints at a context-specific challenge.

Among the farmers registered for our service, it is notable that 21.9% intentionally asked their telephone service providers to unsubscribe from mass SMS campaigns during the service.

Furthermore, the recent passage of the Kenyan Data Protection Act demonstrates the Kenyan government's and its citizens' awareness of the need to regulate mobile-based interventions.

While we do not believe these elements alone explain the lack of significant results, they are interesting alternative mechanisms that could help understand some deviation from other evaluations.

CONCLUSION

While the DEAS program was positively received by many farmers, with a favorable NPS® of 6.32, our causal analysis reveals that the program did not achieve its primary goal of increasing farmers' knowledge of good agricultural practices. This study aimed to replicate the success seen in Rwanda’s coffee sector in the macadamia sector of Kenya. Despite the promising perceptions from beneficiaries, neither the RCT nor the Difference-in-Differences analysis found significant effect on learning outcomes. These findings highlight the need to distinguish between positive perceptions and the actual impact of services, enabling organizations like Root Capital to make informed decisions about program effectiveness.

The evaluation process also highlighted important challenges that warrant attention. Issues such as attrition, low response rates, and limited take-up of the service were encountered during service implementation. Many farmers did not recall receiving advisory service messages. Acknowledging these challenges is crucial as we aim to refine and enhance the program in the future. Overall, while the initial feedback indicates that DEAS delivered value, further efforts are needed to ensure that all farmers can fully benefit from the service.

DISCUSSION

Reporting null results in our evaluations is always challenging. While negative findings can be equally informative, they often face publication bias, leading to a skewed understanding of the field. Furthermore, donors or our programmatic teams might feel disappointed thinking their efforts or support was not fruitful despite significant investments of time and resources. However, this evaluation informs our programmatic decisions on how to move forward with the DEAS service. While reaching out to farmers through SMS or IVR methods can be positive for building rapport and loyalty, it is clear that, in their current state, these methods are not effective tools for knowledge transmission in the Kenyan macadamia industry. We believe that the evaluation also highlights the importance of robustly assessing the impact of services and programs beyond relying solely on perception-based metrics

The contrast between farmers' positive survey responses and their actual performance on knowledge assessments reinforces the importance of rigorously evaluating service impacts beyond just perceived improvements. In recent years, there has been a growing reliance on perception-based measures to claim improvements, quantify impact, or compare services using benchmarks drawn from perception surveys. While these methods offer valuable insights into how services are received and how firms are perceived relative to their peers, it is essential to complement them with objective measures of the intended program outcomes. Although DEAS was well-received by survey respondents, the rigorous evaluation indicates that it fell short of its goal to improve farmers' knowledge of good agricultural practices.

TECHNICAL APPENDIX

Regression Model and Results

This evaluation intended to estimate the attributable impact of DEAS on farmers knowledge of GAPs that improve productivity, sustainability, and quality of macadamia nut. Given that in the past Root Capital had assessed the effect of DEAS in conjunction with other advisory services, we aimed to use an RCT design to isolate the effect the SMS or IVR intervention had service participants. Given the successful randomization of the messaging campaign, it is possible to assume the only difference between farmers of treatment and control groups was receiving or not all the content delivered through the SMS messages.

Our main objective was estimating the effect of DEAS on a Knowledge Score (KS) for good agricultural practices (tailored in different ways to recognize the content different treatment arms were exposed to) or on individual questions of the endline survey. The effect of DEAS (β_1) can be estimated through the following econometric model:

$$Y_{i,c} = \beta_0 + \beta_1 Treatment_{i,c} + \epsilon_{ic}$$

Where $Y_{i,c}$ represent the outcome of interest in of individual i in county c after the service conclude, $Treatment_i$ is a binary variable equal to 1 if the farmer received the intervention or 0 if they received only partial information. As being assigned to the treatment (e.g. receiving the SMS message) does not necessarily imply the farmer appropriated the full treatment content, we interpret these coefficients as ITT parameters. We also cluster robust standard errors at the county level. For ease of interpretation and transparency we used this simple OLS estimation through the paper, finding no statistically significant effect of our intervention in any learning outcomes. However, we are convinced that the adequate estimation strategy should be using the LATE given that only 61.4 % of the 546 respondents who completed the endline survey actually remembered receiving the training material.

Compliance, the Definition of Take-up, and Alternative Estimation of the Effects (ITT vs. LATE) and Correlates of Take-up

As mentioned before, we think the correct estimation strategy to accurately represent the effect of DEAS should be the LATE, using the take-up question as the instrumental variable. The ITT analysis for all key outcomes indicates that, on average, those who were assigned to receive all training material were not more likely to have higher Knowledge Scores, and the effect was not statistically significant even after pooling different treatment groups for comparison purposes. Using the take-up question as instruments, also shows no statistically significant effects on knowledge scores as shown in table 6. The change in the coefficients is mostly explained because the instrumental variable is a weak instrument to explain the assignment to receive DEAS. It is important to note with lower rates of compliance the minimum detectable effect assumption might also be compromised, but in this evaluation, we do not believe we are under powered to find significant effects of the service.

Table 6: Regression Results – RCT / ITT & RCT / LATE

Variables	Overall Knowledge Score (KS) - Q1 -Q9			KS Q4 - Q6	KS Q7 -Q9
	(1) G0 vs G1&G2	(2) G0 vs G1	(3) G0 vs G2	(4) G0 & G1 vs G2	(5) G0 & G2 vs G1
a) RCT - ITT					
DEAS Intervention	**-0.0147**	**-0.0252**	**-0.00417**	**0.0118**	**-0.000996**
	(0.0143)	**(0.0153)**	**(0.0148)**	**(0.0224)**	**(0.0148)**
Constant	0.718***	0.729***	0.707***	0.769***	0.853***
	(0.0119)	(0.0112)	(0.0145)	(0.0142)	(0.0122)
Observations	1,120	737	734	739	638

Asterisks denote statistical significance at the confidence interval of 99% (***), 95% (**) and 90% (*), respectively.

Variables	(1) G0 vs G1&G2	(2) G0 vs G1	(3) G0 vs G2	(4) G0 & G1 vs G2	(5) G0 & G2 vs G1
b) RCT - LATE					
DEAS Intervention	-36.24	-3.935	7.412	5.205	-3.737
	(791.5)	(13.16)	(33.85)	(8.570)	(6.053)
Constant	11.87	2.637	-2.680	-2.572	3.339
	(242.9)	(6.290)	(15.67)	(5.568)	(4.001)
Observations	532	344	351	532	532

Recognizing the importance of robustness checks for our estimation strategy, the evaluation design incorporated a baseline survey. This survey served to verify the effectiveness of randomization prior to DEAS implementation and also to create baseline data for potential use in a Difference-in-Differences analysis. However, due to the small number of participants with both baseline and endline survey responses, we were unable to reliably estimate treatment effects using this method.

Finally, given the null effects resulting from the evaluation we did not further explore other variables that could help explain correlated of take-up. For example, Aker Collier, and Vicente (2017) found that women are less likely to report receiving and reading messages (when comparing self-reported responses to actual delivery data). Rounseville, Oviedo and Vera (2019) found that education and mobile reception are positively correlated with message adoption. However, female-headed households were less likely to be correlated with receiving messages. Given these evidence, future evaluations of DEAS should take into account no only indicators for take-up but also control for these other variables that correlate to take-up.

REFERENCES

Aker, Jenny, Girija Bahety, Megan Rounseville, Melita Sawyer, and Juan C. Taborda. 2021. *Providing Information via Technology: Evidence from SMS Interventions*. *Working Paper*.

Aker, Jenny C., Paul Collier, and Pedro C. Vicente. 2017. "Is Information Power? Using Mobile Phones and Free Newspapers during an Election in Mozambique." *The Review of Economics and Statistics* 99(2):185–200. doi: 10.1162/REST_a_00611.

Aker, Jenny C., Ishita Ghosh, and Jenna Burrell. 2016. "The Promise (and Pitfalls) of ICT for Agriculture Initiatives." *Agricultural Economics* 47(S1):35–48. doi: 10.1111/agec.12301.

Bain & Company Inc. 2017. "Measuring Your Net Promoter Score." *Bain*. Retrieved September 10, 2024 (https://www.netpromotersystem.com/about/measuring-your-net-promoter-score/).

Naeve, Katie, Sabrina Kerin, Erick Sakwa, and Isabel DoCampo. 2022. "Evaluation of a Mobile-Based Agronomic Extension Services Program for Coffee Farmers in Rwanda."

Rounseville, Megan, Ana Oviedo, and Priscila Vera. 2019. ""Can Text Messages Improve Children's Health? A Case from Rural Ecuador."" Forthcoming.

Sauro, Jeff. 2017. "Changing the Net Promoter Scale: How Much Does It Matter?" *MeasuringU*. Retrieved September 10, 2024 (https://measuringu.com/nps-scale-change/).

Taborda, Juan C., and Pablo Busto. 2023. "Evaluation of a Mobile-Based Agronomic Extension Services Program for Coffee Farmers in Rwanda – Phase 2."