Assessing the Net Economic Impact of the Farmer Field and Business Schools

A Micro-econometric Analysis of the Benefits and the Costs of the FFBS Program in Kenya between 2016 and 2022

giz Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH

Assessing the Net Economic Impact of the Farmer Field and Business Schools

A Micro-econometric Analysis of the Benefits and the Costs of the FFBS Program in Kenya between 2016 and 2022

Contents

Appendices

1 Introduction

The Global Project (GP) "Promotion of nutrition-sensitive potato value chains in East Africa" is implemented across three fields of action, concerned with 1) improving the quality and quantity of potato production and marketing, 2) improving nutritional habits and basic hygiene, and 3) improving coordination within the potato sector. Across these three fields of action, the GP will be evaluated in the course of a Central Project Evaluation (CPE), which will assess the project along the six DAC-OECD evaluation criteria relevance, coherence, effectiveness, efficiency, impact and sustainability. As part of the Kenya country component evaluation, the Farmer Field and Business School (FFBS) approach of the GP was to be evaluated quantitatively.

The FFBS offers practical training on the Good Agricultural Practices (GAP) and farm business management in the context of potato farming, with some schools (termed "nutrition-integrated") additionally offering teaching on dietary diversity, hygiene, kitchen garden food production, preparation, and preservation, and child nutrition. An FFBS typically consists of a trainer and 15 to 25 participant farmers who come together on a demonstration plot for 15 day-long sessions occurring over the course of a season. Between 2017 and 2022, 991 FFBS have been implemented, of which 408 were led by county staff and 583 by lead farmers. Implementation began in Nyandarua and Bungoma, and was extended to Elgeyo Marakwet and Trans Nzoia in 2019. In total, 17.927 farmers have been trained.

The quantitative evaluation of the FFBS program has been carried out; it is documented in its entirety in the report at hand. In the course of the evaluation, a questionnaire has been designed and administered to a sample of potato farmers that has been carefully stratified along multiple dimensions. A matching model has been developed to allow a statistically valid comparison between FFBS-trained and non-FFBS-trained farmers in the sample; additionally, inferential models have been built to be able to cleanly estimate the effect of the FFBS on ${}^{Yield}/_{ha}$ and ${}^{Price}/_{kg}$ of potatoes harvested by the beneficiaries, allowing us to understand the impact of the FFBS on revenue, cost and surplus. Finally, program expenditures have been carefully tallied, so that development and implementation costs of the FFBS may be derived.

These calculations have been undertaken to arrive at a single figure: the net economic impact of the FFBS. The net economic impact represents an estimate for the total "social surplus" generated by the FFBS: After program costs, how much net value, expressed in KES, has been created by the program in the hands of its beneficiaries? Alternatively, this figure will be stated as a benefit-cost ratio: What is the social value created by each unit of currency invested in the FFBS program?

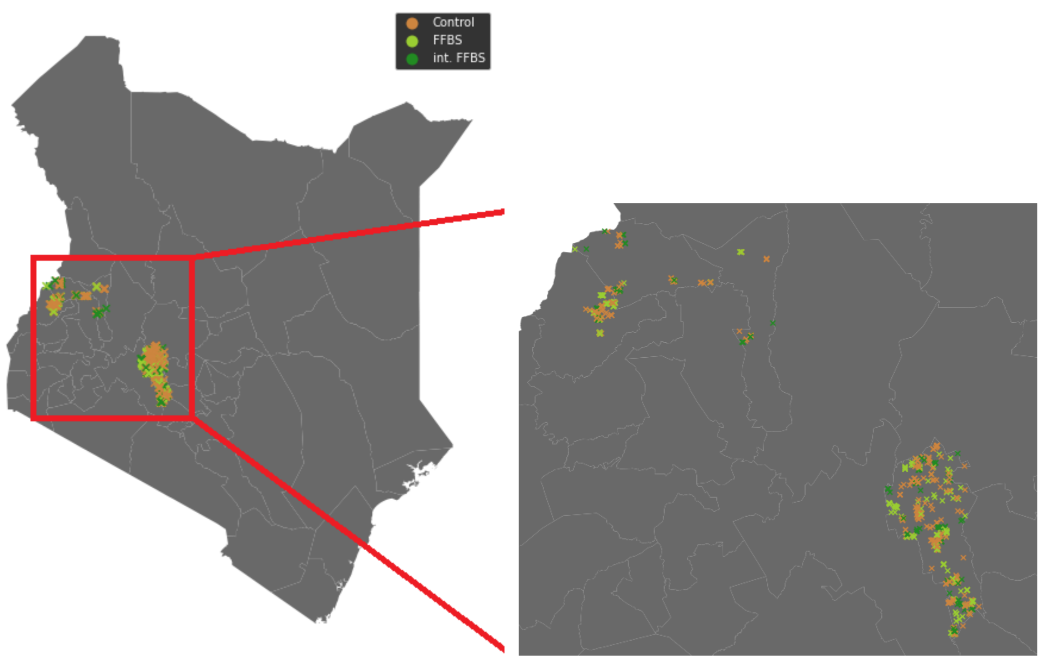

Figure 1: The Geographical Distribution of Treatment and Control

2 Data

The net economic impact figure that we aim to supply is an empirical result; that is, it is derived (almost) entirely from collected data. This data was collected to feed a set of models that would, in proper arrangement, yield a statistically valid net economic impact figure. The arrangement of statistical models will be outlined in later sections; in this section, we focus on the data at hand, without which this analysis would be impossible. We will relate the shape and size of our dataset, the targeted sampling of FFBS-trained and non-FFBS-trained farmers, and the scheme by which we match these two sample groups to allow a statistically valid comparison.

2.1 Sampling

The data was collected as part of a survey among potato farmers conducted in October 2022. The questionnaire covers up to 291 items pertaining to demographics, land ownership and use, agricultural practices and their related costs, tuber yields and prices fetched, and nutritional preferences and practice. Questions related to the agronomic process were asked in reference to the 2022A season, which ended in August 2022. In total, the dataset consists of 752 respondents.[1] Of these, 387 are not FFBS-trained: these constitute our control group. 365 farmers are FFBS-trained: these constitute our treatment group. The treatment group is further subdivided into two groups. The sample of "pure" FFBS-trained farmers, who received training only on the core curriculum of GAP and farm business management, will be denoted simply *FFBS*. The sample of "nutrition-integrated" FFBS-trained farmers, who in addition to the core curriculum received nutrition-related training, will be denoted *int. FFBS*. Both samples together will be denoted *Treatment*. The sample of non-FFBS farmers will be denoted *Control*. All farmers in the *Treatment* were trained in or before Season 2021A, which ended in August 2021, thus have had at least two full seasons to put the learned teachings into practice. The geographical distribution of sampled farmers is displayed in Figure 1: On this map, each cross represents a single respondent, with color representing the corresponding sampling group. Aggregated to county-level, the resulting sample sizes across sampling groups are outlined in Table 1.

	Control	FFBS	int. FFBS	
Nyandarua	279	181	82	**542**
Bungoma	59	46	10	**115**
Trans Nzoia	32	9	22	**63**
Elgeyo Marakwet	17	7	8	**32**
	387	**243**	**122**	**752**

Table 1: Number of Observations across Treatment and County

The sampling process was carried out in stratified manner, across sample groups and counties. For the *Treatment*, this means that the share of farmers sampled in (for example) Nyandarua relative to all *Treatment* farmers should reflect the proportion of farmers trained by the FFBS in Nyandarua, relative to all farmers trained. Similarly, the proportion of farmers in the *int. FFBS* sample should reflect the proportion of farmers that received the nutrition-integrated FFBS training, versus all farmers that received training, be it pure or nutrition-integrated.[2]

The size of *Control* was optimized using data collected in 2018 and analysed in a previous impact assessment (Vagliano, 2019). The optimization approach taken was to maximise the statistical power of a test for treatment effect of FFBS training on $^{Yields}/_{ha}$ given a budget available for surveying. The resulting control-to-treatment sampling target ratio was established to be 1.1, that is, that there should be 11 *Control* group members for 10 *Treatment* group members. Methodologically, this procedure is explained in depth in Appendix A.

[1] 758 farmers were surveyed; however, six farmers declined to participate after agreeing to an appointment with the enumerator initially.

[2] Note that within the scope of stratification, "all" farmers more precisely denotes "all farmers trained in or before Season 2021A". These constitute the relevant base population of 14,028 farmers, as opposed to the entire population of trained farmers, which consists of 17,927 farmers.

Table 2 outlines the number of trained farmers for each county-sample group combination. Given proportional representation, and a control-to-treatment sampling rate of 1.1, the theoretically optimal representative sample comprising 752 farmers is outlined. This figure is contrasted with the actual sample achieved, to give the difference between plan and outcome. Overall, the sampling outcome is highly satisfactory in every respect. The aggregate sizes of sample groups *Control*, *FFBS*, and *int. FFBS* are within 5% of target. The targeted control-to-treatment ratio of 1.1 was undershot slightly, standing at 1.06. Deviations within the county representation are relatively small also, with the exception of Trans Nzoia, which is moderately underrepresented.

		Control	FFBS	int. FFBS	
Trained Farmers	Nyandarua		6,936	2,841	
	Bungoma		1,857	441	
	Trans Nzoia		84	1,330	
	Elgeyo Marakwet		277	262	
Optimal Representative Sample	Nyandarua	275	177	72	
	Bungoma	65	47	11	
	Trans Nzoia	40	2	34	
	Elgeyo Marakwet	15	7	7	
Actual Sample	Nyandarua	279	181	82	
	Bungoma	59	46	10	
	Trans Nzoia	32	9	22	
	Elgeyo Marakwet	17	7	8	
Difference	Nyandarua	4	4	10	**18**
	Bungoma	-6	-1	-1	**-8**
	Trans Nzoia	-8	7	-12	**-13**
	Elgeyo Marakwet	2	0	1	**3**
		-8	**10**	**-2**	

Table 2: Deviation from Planned Stratified Sampling

2.2 Matching

For any given variable of interest, comparisons between a *Control* and a *Treatment* group can serve to understand the difference between a typical FFBS-trained and a typical non-FFBS farmer. However, we cannot as easily find the *causal* effect of the FFBS on the variable in question. This is because the typical FFBS-trained farmer might differ from the typical non-FFBS trained farmer in more ways than just training received. For example, assume we are interested in the effect of the FFBS on $^{Yield}/_{ha}$, but we suspect that the typical farmer who decided to participate in FFBS training is more motivated than the typical non-FFBS trained farmer. If we now simply compare the average $^{Yield}/_{ha}$ in the *Treatment* with the average $^{Yield}/_{ha}$ in the *Control* group, the difference in yields might not only be driven by participation in the FFBS, but also reflect the effect of higher motivation. In this case, our naive measurement of the effect of the FFBS, the difference between sample averages within *Treatment* and *Control*, would be confounded and useless. As we will see in brief, the fundamental cause of this issue is selection bias.

In a perfect world, the researcher could observe the level of the variable of interest given that a farmer i is trained, and the level given that that same farmer i is not trained by the FFBS. This way, the only factor that would differ is FFBS training, and all other confounding factors, among which the degree of motivation, would be kept constant. In such a setting, the researcher could estimate a treatment effect, which, for any given member i of the population, is calculated as the difference in the level of the variable between the situation in which i is assigned to *Treatment*, and the situation in which i is assigned to *Control*. The average treatment effect (ATE) is simply the average of all these differences across the entire population; the average treatment effect among the treated (ATT) is the average of all these differences across the population of the treated. If the variable is chosen to be the $^{Yield}/_{ha}$, for example, the ATE would indicate the average increase in $^{Yield}/_{ha}$ directly caused by the FFBS if every potato farmer in Nyandarua, Bungoma, Elgeyo Marakwet and Trans Nzoia was trained; the ATT would indicate the same increase only for those actually trained. Naturally, since a given farmer is assigned either to *Treatment* or *Control*, but never both, the *individual* treatment effects are unobservable; however, the *average* treatment effects ATE and ATT can be easily estimated without bias in a sample where selection into either *Treatment* or *Control* is randomised.[3]

In the case of our study, however, selection into the sample is not randomized. Instead, farmers self-select into participating in the FFBS, or they select out of it, either explicitly or implicitly. This introduces selection bias, which poses a major methodological hurdle to estimating the ATE or the ATT. To circumvent this hurdle, we implement Propensity Score Matching (PSM), which allows us to statistically approximate a situation in which selection into the sample is randomized, so that treatment effects can again be calculated (Rosenbaum & Rubin, 1983). In effect, PSM estimates for each respondent the probability that said respondent selects into the *Treatment*. *Control* group member outcomes are then weighted: The higher the probability, the higher the weight. In this way, PSM creates a *Matched Control* group which is similar to, and therefore comparable with, the *Treatment* (Austin, 2016). If the propensity score estimate is adequate, the difference between the average outcomes across these two groups (which will in the following often be reported) are then unbiased estimators for the ATT. However, since no matching model is perfect, we can not guarantee all selection bias to be eliminated. Hence, as an additional safeguard against selection bias, we will further limit the choice of variables that we compare *Treatment* and *Matched Control* on. This additional safeguard will be explained later, in Section 4.1.

To implement PSM, a model has to be estimated. This model aggregates farmer characteristics into a single number: the probability of selection into *Treatment*. These probabilities are then used to assign weights to each member of the *Control* group. These weights are selected by the model so that the weighted *Control* group on aggregate corresponds very closely to the *Treatment* sample, with respect to the farmer characteristics that drive selection into the FFBS. It is in this sense that the weighted *Control* group is termed *Matched Control*: A typical member of *Matched Control* matches the characteristics of a member of *Treatment* member as well as possible. In this way, selection bias along these characteristics is reduced. The characteristics which enter our model are demographic, or pertain to the attitude of a farmer towards various factors. Table 3 outlines these demographics and attitudes. With the exception of those rows marked with an asterisk, all factors enter into the propensity score model. It can be well observed that the *Matched Control* is indeed more similar to the *Treatment* than the *Control* group, across most model constituents.

The methodology of PSM is outlined in more depth in Appendix B, along with the rationale behind the inclusion of certain characteristics. The regression parameters of the propensity score model are outlined in Table A8 of Appendix G.

3 Descriptive Analysis of Matched Samples

In this section, we will apply the matching methodology to compare a typical FFBS-trained farmer and his "similar", non-FFBS-trained counterpart with respect to some key variables. Of particular interest are increases in GAP adoption between *Treatment* and *Matched Control*, which we will relate in Section 3.1. We will also take a comparative look at agronomic variables such as $^{Yield}/_{ha}$, $^{Price}/_{kg}$, revenues and costs in Section 3.2.

3.1 GAP Adoption

Table 4 outlines the proportion of farmers correctly adopting a given GAP. It also indicates in bold the difference between *Treatment* and *Matched Control*, which, given a perfect propensity score model, is an unbiased estimate for an average treatment effect among the treated (ATT), that is, the extent to which the increase in GAP adoption among trained farmers can be causally attributed to the FFBS. Next to these differences, stars indicate to what extent these estimates for the ATT are statistically significant, with respective significance levels α reported in the

[3] In fact, in this case, the ATE and the ATT are equal in expectation.

		Control	Matched Control	Treatment
Personal Info	Age	47	51	51
	Experience in Potato Production	10	11	11
	Female	55%	67%	63%
	Male	45%	33%	37%
Education	No Schooling	2%	2%	2%
	Primary	43%	46%	45%
	Secondary	44%	42%	44%
	University	11%	9%	7%
	Adult Education	0%	0%	1%
Land Ownership	Total Land (in hectares)	1.08	1.11	1.06
	*... of which: potato**	0.31	0.31	0.33
	*... of which: potato, leased**	0.11	0.10	0.13
	*... of which: potato, main plot**	0.29	0.29	0.29
Source of Agronomic Info	Input suppliers	26%	25%	27%
	Extension services*	23%	32%	89%
	Organised farmer groups*	9%	10%	56%
	Media	58%	64%	59%
	Informally, through peers	67%	36%	37%
Interest in Learning About Agronomics	Not interested at all	1%	0%	1%
	Mostly uninterested	6%	3%	3%
	Interested	55%	54%	48%
	Very interested	39%	44%	48%
Perceived Demand for Ware Tubers	I struggle to sell what I harvest	16%	8%	6%
	I can sell what I harvest	13%	9%	15%
	I could sell a bit more than I harvest	20%	21%	16%
	I could sell much more than I harvest	51%	62%	63%
Preference for Food Diversity	I am content eating the same foods every day	3%	1%	1%
	I need a little variety once in a while	27%	21%	13%
	I need some variety in my diet regularly	52%	53%	63%
	I need a lot of variety in my diet every day	16%	25%	22%

Rows marked with an asterisk (*) indicate variables which were not used to match *Control* and *Treatment*.

Table 3: Respondent Demographic Data

notes to Table 4. A strong majority of GAP see a statistically significant increase in adoption between *Matched Control* and *Treatment*. On average,[4] GAP adoption stands at 51% in *Control*, at 53% in *Matched Control* (translating to roughly 13 GAP adopted out of 25, on average, in both groups), and at 71% in *Treatment* ($\approx$ 18 out of 25). In this section, we will point out the most notable and significant increases in GAP component adoption attributable to the FFBS.

3.1.1 Seed Selection

Two thirds of trained farmers use selected, clean or certified tubers, a figure significantly larger than in *Matched Control*, where only one in eight farmers plant selected, clean or certified seed. The estimated ATT for sprouting stands at a moderate, yet highly significant, 13 percentage points. As for seed quantity, there are slight improvements: 41% of FFBS-trained farmers meet the recommended 800 to 1,000 $^{kg}/_{acre}$, as opposed to 34% in *Matched Control*. As we will see in Section 5.1.1, seed quantity typically falls below the requirements and model optima for both *Treatment* farmers and, to a larger extent, *Matched Control* farmers.

3.1.2 Soil Fertility Management

Soil testing is done in about one in three cases in *Treatment*, and almost never done in *Matched Control*. Almost all farmers in *Treatment* use the right type of fertilizer, up from already high levels in *Matched Control*. The proportion of farmers using the correct quantity of fertilizer

[4]The average here denotes a weighted mean across all GAP, with weights of $^1/_2$ assigned to the a and b components of GAP 17, 18 and 19, and weights of 1 otherwise.

	#	Practice	Description	Control	Matched Control	Treatment	Average Treatment Effect among the Treated (ATT)
Seed Selection	1	Seed Quality	Seed tubers used are either positively selected, clean (i.e., grown from certified seed) or certified.	13%	15%	66%	**+51%** ***
	2	Sprouting	Potatoes are sprouted to any length.	59%	66%	79%	**+13%** ***
	3	Seed Quantity	Seed density lies between 800 and 1000 $^{kg}/_{acre}$.	32%	34%	41%	**+7%** *
Soil Fertility Management	4	Soil Testing	Soil testing and analysis is conducted.	4%	6%	31%	**+25%** ***
	5	Fertilizer Type	Soil test recommendations are fully applied, or, alternatively, the farmer applies organic fertilizer or inorganic fertilizer containing phosphorus.	82%	88%	96%	**+8%** ***
	6	Fertilizer Quantity	Soil test recommendations are fully applied, or, alternatively, the recommended density (for manure: 2000-4000 $^{kg}/_{acre}$; for fertilizer: varies by product, margin of error of 25 $^{kg}/_{acre}$ allowed) is applied.	12%	15%	41%	**+26%** ***
Land Preparation	7	Timing	Land preparation is done before the onset of rains.	76%	80%	90%	**+10%** ***
	8	Ploughing Depth	When using a tractor or an ox-drawn plough, the land is plowed to a depth of around 16 to 35 cm. When using a jembe, the land is ploughed to a depth of 11 to 25 cm.	79%	83%	86%	**+3%**
	9	Steps	For old, fallow or virgin land, at least two, three or four (respectively) of the following four steps have to be taken: 1) clearing (mechanically or with herbicide), 2) primary ploughing or chiseling, 3) secondary ploughing, 4) furrowing or harrowing.	65%	62%	67%	**+5%**
Planting Practices	10	Seed Spacing	Potatoes are planted with a distance of 30 cm between potatoes within rows, and a distance of 75 cm between rows.	27%	33%	80%	**+47%** ***
	11	Planting Depth	Potatoes are planted at a depth of 16 to 20 cm.	13%	16%	16%	**+0%**
Weeding, Hilling & Thinning	12	Frequency	Weeding and Hilling is done at least twice.	39%	42%	60%	**+18%** ***
	13	Height of Hills	The height of hills is between 20 and 30 cm.	51%	56%	62%	**+6%** *
	14	Thinning	If seed is planted more densely than recommended, potatoes are checked for plant density, and thinning is done as necessary.	38%	40%	84%	**+44%** ***
Pest & Disease Management	15	Scouting	Potatoes are scouted for pest and disease.	69%	73%	98%	**+25%** ***
	16a	Pest Control: Measures	As in 17a), but some manual, biological or chemical measures are taken to combat the pest.	66%	61%	84%	**+23%** ***
	17a	Pest Control: Toxicity	When a farmer's crop is affected by pest, the farmer does not use toxic chemicals (WHO Hazard Classes Ia and Ib).	100%	100%	98%	**-2%**
	18a	Pest Control: Quantity	As in 16a), but the dilution of chemical used (measured in $^{mL}/_{L}$ or $^{g}/_{L}$) is compliant with the dilution recommended by the manufacturer.	57%	56%	72%	**+16%** ***
	16b	Disease Control: Measures	As in 17b), but some manual, biological or chemical measures are taken to combat the disease.	84%	84%	98%	**+14%** ***
	17b	Disease Control: Toxicity	When a farmer's crop is affected by disease, the farmer does not use toxic chemicals (WHO Hazard Classes Ia and Ib).	100%	100%	100%	**-0%**
	18b	Disease Control: Quantity	As in 16b), but the dilution of chemical used (measured in $^{mL}/_{L}$ or $^{g}/_{L}$) is compliant with the dilution recommended by the manufacturer.	66%	68%	64%	**-5%**
	19	Chemicals: Storage	If chemicals are present, they are stored in a dedicated place of storage away from the main house.	71%	73%	81%	**+8%** **
	20	Chemicals: Protection	If pesticides are applied, three or more protective items are worn among the following: coat, goggles, nose mask, gumboots, gloves, helmet, old clothes, face mask.	52%	49%	77%	**+28%** ***
Crop Rotation & Fallowing	21	Non-Consecutive Potato	Potato was not planted in the same field in which potatoes were planted in the current season.	82%	87%	95%	**+9%** ***
	22	Recommended Rotation	The crops planted before, during and after the season in question all come from a different family of plants.	44%	51%	67%	**+17%** ***
Harvesting & Post-Harvesting Practices	23	Dehaulming	Potatoes are dehaulmed before harvesting.	12%	15%	64%	**+49%** ***
	24	Harvesting when Mature	Depending on the variety of potato planted, tubers are harvested a given number of days after planting, allowing for a margin of 15 days.	32%	32%	30%	**-2%**
	25	Sorting and Grading	Potatoes are sorted and graded after harvest.	86%	87%	96%	**+10%** ***
				51% ($\approx$ $^{13}/_{25}$)	**53%** ($\approx$ $^{13}/_{25}$)	**71%** ($\approx$ $^{18}/_{25}$)	**+17%** *** ($\approx$ $+^{4}/_{25}$) ***

Difference between *Treatment* and *Matched Control* highlighted.
$*: p < 0.05$, $**: p < 0.01$, $***: p < 0.001$

Table 4: The Adoption of the Good Agricultural Practices

increases to around 40%, up from 15%. In Section 5.1.2, we will outline that the optimal quantity of planting fertilizer is undershot considerably on average, while the optimal quantity of top dressing fertilizer is overshot slightly.

3.1.3 Land Preparation

In the domain of land preparation, increases in adoption are more subdued. The only increase in adoption which can be attributed to the FFBS with statistical significance is the increase in the probability to prepare the land before the onset of rains, which stands at 10 percentage points. For both ploughing depth and the land preparation steps, no significant difference between *Matched Control* and *Treatment* can be established.

3.1.4 Planting Practices

The share of farmers observing inter-plant and inter-row distances of 30 and 75 cm respectively increases by a staggering 47%, from 33% in *Matched Control* to 80% in *Treatment*. The correct depth of planting (between 16 and 20 cm) is scarcely observed throughout the sample.

3.1.5 Weeding, Hilling & Thinning

In all GAP associated with Weeding, Hilling and Thinning, moderate differences between *Matched Control* and *Treatment* are recorded. The share of farmers doing at least two rounds of weeding and hilling increases significantly, from around 40% in *Matched Control* to around 60% in *Treatment*. If seed density is higher than recommended, most *Treatment* farmers thin their crop, whereas in *Matched Control*, this practice is reported only by a minority of respondents. The observance of recommended hill heights is higher in FFBS-trained farmers than in similar non-FFBS farmers, but only at a significance level of 5%.

3.1.6 Pest & Disease Management

An FFBS-trained farmer is almost certain to scout for pest and disease, compared to a similar, but non-FFBS-trained farmer, who scouts in 73% of cases, on average. With respect to pesticide and fungicide toxicity, virtually all farmers in the entire sample apply chemicals not designated as extremely or highly hazardous by the WHO. However, *Treatment* farmers are moderately more likely to combat pests and diseases using appropriate measures. The adherence to safety measures with respect to the handling of pesticides is significantly higher in *Treatment* than in *Matched Control*. Chemicals are more likely to be stored in a dedicated place away from the house, and adequate protective equipment is worn more frequently.

3.1.7 Crop Rotation

Significant increases in the adherence to crop rotation GAPs can be attributed to the FFBS, with almost all farmers in *Treatment* not planting potato consecutively, and around two thirds of farmers planting three different families of crops before, during and after the season in question.

3.1.8 Harvesting & Post-Harvest Practices

Before harvest, almost two thirds of trained farmers dehaulm their potatoes, highly significantly up from a lowly 15%. After harvest, almost all *Treatment* farmers sort and grade their potatoes, up from already high levels in *Matched Control*.

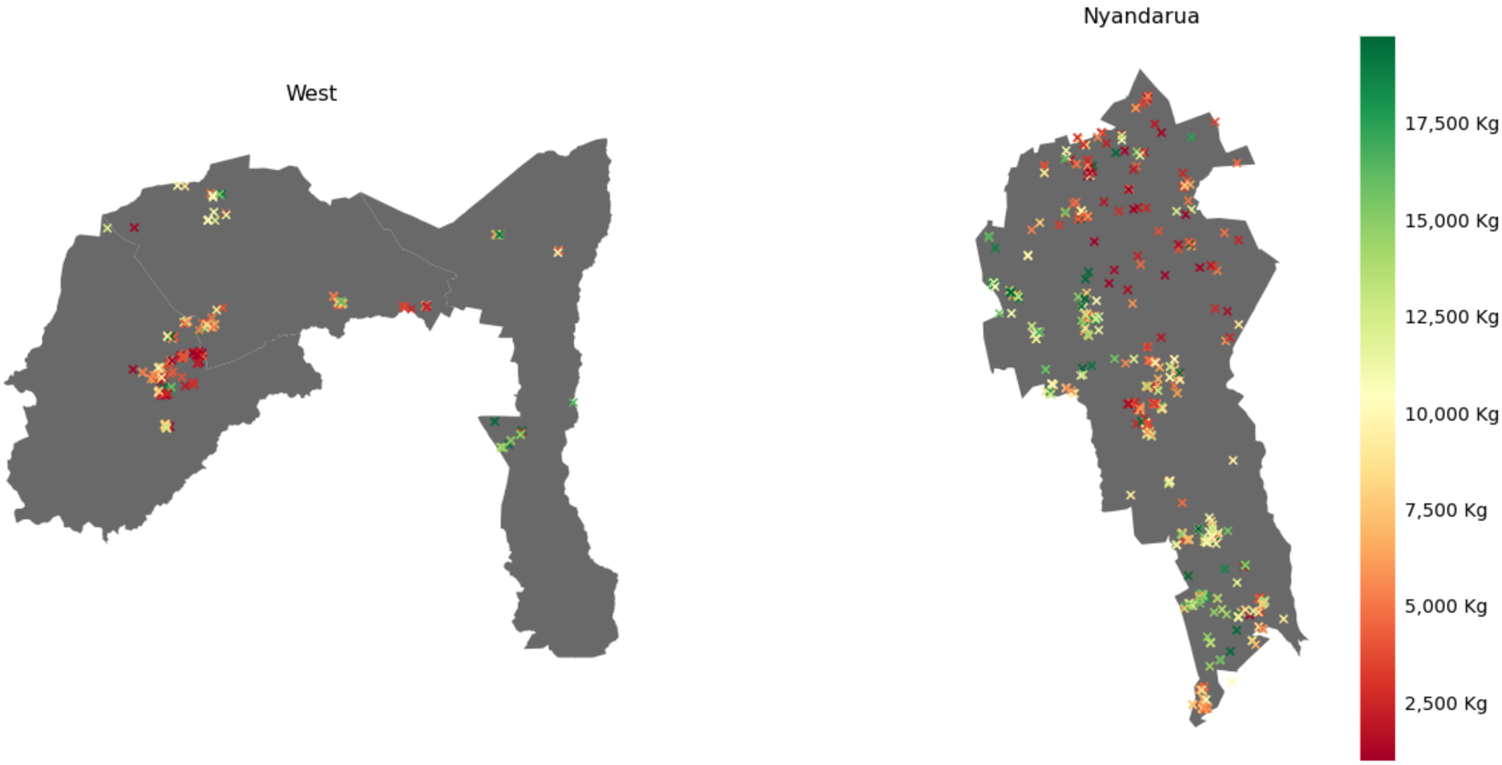

Figure 2: The Geographical Distribution of $^{Yield}/_{ha}$

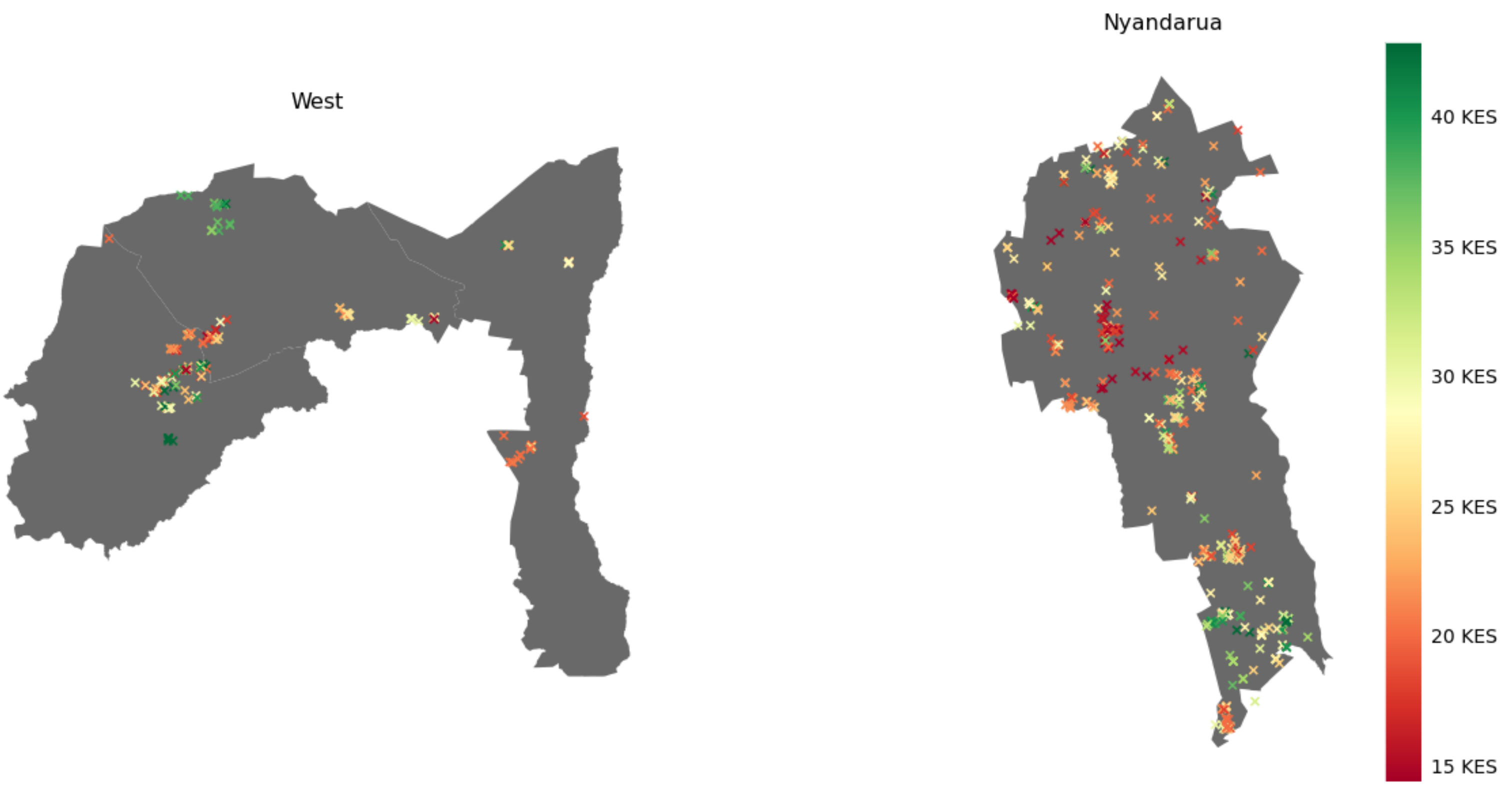

Figure 3: The Geographical Distribution of $^{Price}/_{kg}$

3.2 Agronomics

Figures 2 and 3 show the geographical distribution of $^{Yield}/_{ha}$ and $^{Price}/_{kg}$ respectively. In the parts of Bungoma from which respondents were sampled, $^{Yield}/_{ha}$ is very low. In Nyandarua, $^{Yield}/_{ha}$ is elevated in the west and the south of the county, whereas lower figures are observed in the north-eastern part of the county. A point of interest was to inferentially assess whether prices are responsive to over- or undersupply of tubers, as proxied by $^{Yield}/_{ha}$. At first glance, there is some indication that this might be the case: ware prices seem elevated in Bungoma (where $^{Yield}/_{ha}$ is low), and lower in western Nyandarua (where $^{Yield}/_{ha}$ is higher relative to the rest of the county). However, as we will see later, a price response to quantities supplied could not be confirmed.[5]

For a county-level comparison between our sample groups, consider Table 5, which shows yields and prices across the sample. Overall, a *Control* group farmer harvests on average 7.4$^{tons}/_{ha}$, compared to 10.1$^{tons}/_{ha}$ in the *Treatment*. However, some of this difference is explained by self-selection: The *Matched Control* exhibits higher yields than the pure *Control* by about 0.5$^{tons}/_{ha}$. Nevertheless, the comparison between FFBS-trained farmers and similar farmers which are not trained by the FFBS still reveals a staggering difference in $^{Yields}/_{ha}$ of more than 2$^{tons}/_{ha}$. Ware potato prices fetched by trained farmers are significantly higher than for similar, non-FFBS farmers in Nyandarua. In the three western counties, differences are negative, but insignificant. In general, sample sizes for these counties are individually small. Hence, in Table 6 below, western counties are aggregated into one category. The following paragraphs outline the value created by potato yields given a tuber's potential different uses, and the different prices (or, more generally, unit values) fetched by each of these uses, before contrasting these with the costs associated with agronomic production, to finally arrive at an economic surplus figure generated on average per hectare per season.

		Control	Matched Control	Treatment	
Yield/ha *(in Kg)*	Bungoma	5,918	5,429	5,829	+7%
	Elgeyo Marakwet	11,865	9,501	14,945	+57% **
	Nyandarua	7,464	8,273	10,721	+30% ***
	Trans Nzoia	7,196	7,869	10,667	+36% *
	All	**7,400**	**7,857**	**10,140**	**+29%** ***
Price/kg *(in KES)*	Bungoma	35	32	29	-9%
	Elgeyo Marakwet	24	26	23	-12%
	Nyandarua	24	23	29	+26% ***
	Trans Nzoia	27	30	27	-10%
	All	**25**	**25**	**28**	**+12%** ***

Difference between *Treatment* and *Matched Control* highlighted.
$* : p < 0.05$, $** : p < 0.01$, $*** : p < 0.001$

Table 5: Yields and Prices

3.2.1 The Many Uses of A Tuber

After harvest, a tuber can derive value in many ways. Most commonly, a tuber is sold, either as ware or seed. As can be seen in Table 6, higher yields for FFBS farmers translate well into increases in tubers sold as ware or seed. It is also noteworthy that the volume of tubers retained as seed increases everywhere. For trained Nyandaruan farmers, retained seed ($\approx$ 1.8$^{tons}/_{ha}$) approaches the recommended seed rate (2$^{tons}/_{ha}$), indicating that very little seed has to be bought, on average. In the western counties, seed retention is lower, but nevertheless increasing. Despite increases in yields, consumption of harvested potatoes decreases in both Nyandarua and Bungoma between *Matched Control* and *Treatment*. Finally, increases in yields do not seem to correlate with greater storage losses, or more potatoes given away for free.

[5]We tested this hypothesis by adding $^{Yield}/_{ha}$ in the vicinity as a regressor to the model for $^{Price}/_{kg}$ of potatoes sold, which did not result in a significant parameter. The model will be touched upon in Section 4.2, and explained in greater depth in Appendix C.

3.2.2 Value created ...

The economic impact of these differences in yields (and, in Nyandarua, prices) causes a substantial difference in value generated before costs. In Nyandarua, trained farmers generate almost 275,000 KES in economic value per season per hectare, around 110,000 KES more than their untrained counterparts. In the West, the additional value created is more subdued, at around 25,000 KES. The largest part of this difference reaches the beneficiaries in the form of cash: revenues per hectare increase by 100,000 KES in Nyandarua, and 30,000 KES in the western counties. These differences are mostly driven by ware potato sales; in the West, this effect is supplemented by stronger seed potato sales. Value addition, such as the preparation of crisps or potato cakes, does not play a significant role. The economic value of the increased rate of seed retention is reflected in the item *Non-Accounting Value*. This item subsumes all factors from which a farmer derives utility (for example, savings on seed expenses if tubers are retained as seed), but which have no immediate cash impact.[6] These non-accounting items serve to complement the additional revenue created by a trained farmer in Nyandarua, adding another 10,000 KES to the difference in total value created, and slightly offset this difference in the West.

3.2.3 ... costs incurred ...

In Nyandarua, participation in the FFBS correlates with higher costs, driven predominantly by higher seed rates (and, therefore, higher costs of seed),[7] and more intensive soil fertility measures. In the West, costs are slightly lower in *Treatment*, driven by lower labor costs on labor-intensive steps such as planting, weeding and hilling, and harvesting.

3.2.4 ... and profit made!

Netting these costs against value created, the average surplus created per hectare per season can be calculated. In Nyandarua, it amounts to almost $120,000^{KES}/_{ha}$ for *Matched Control* and just over $210,000^{KES}/_{ha}$ for *Treatment*; In the West, surpluses for *Matched Control* and *Treatment* are around $95,000^{KES}/_{ha}$ and $130,000^{KES}/_{ha}$ respectively. In the entire sample, the average surplus stands at $190,000^{KES}/_{ha}$ for *Treatment*, and $110,000^{KES}/_{ha}$ for *Matched Control*.

	Nyandarua			West		
	Control	Matched Control	Treatment	Control	Matched Control	Treatment
Yield$_{/\text{ha}}$ *(in Kg)*	7,464	8,273	10,721	7,233	6,978	8,640
... *of which: sold as ware*	4,194	4,800	7,191	3,998	3,758	5,102
... *of which: sold as seed*	324	413	342	175	111	488
... *of which: kept as seed*	1,502	1,523	1,824	728	850	957
... *of which: consumed*	805	881	745	1,663	1,656	1,503
... *of which: given away*	309	350	383	492	399	387
... *of which: rotten or lost*	330	306	235	178	205	203
Ware Price$_{/\text{kg}}$ *(in KES)*	24	23	29	30	30	27
Seed Price$_{/\text{kg}}$ *(in KES)*	25	24	31	30	30	29
Total Economic Value$_{/\text{ha}}$ *(in KES)*	**161,099**	**173,579**	**282,669**	**179,423**	**185,697**	**212,531**
Revenue$_{/\text{ha}}$	106,000	116,677	217,013	112,619	115,391	144,143
... *of which: sold as ware:*	97,113	106,745	205,204	106,484	111,660	126,559
... *of which: sold as seed:*	8,843	9,888	11,649	4,328	3,169	16,864
... *of which: sold with added value*	44	44	160	1,808	562	719
Non-Accounting Value$_{/\text{ha}}$	55,099	56,902	65,656	66,804	70,306	68,389
... *of which: consumed:*	19,902	21,590	21,311	48,062	48,544	44,777
... *of which: kept as seed:*	35,197	35,312	44,344	18,742	21,762	23,611
Total Economic Cost$_{/\text{ha}}$ *(in KES)*	**54,098**	**55,894**	**70,385**	**98,995**	**92,394**	**82,753**
... *of which: Land Use*	4,961	4,559	5,750	2,666	1,743	2,011
... *of which: Land Preparation*	5,387	5,245	5,960	12,506	10,574	10,308
... *of which: Soil Fertility*	8,040	8,621	14,455	15,433	15,008	14,210
... *of which: Seed*	14,593	15,346	21,643	20,385	18,484	16,466
... *of which: Planting*	3,530	3,348	3,916	8,455	7,679	6,303
... *of which: Weeding and Hilling*	9,681	9,999	9,816	25,914	24,110	19,218
... *of which: Pest and Disease Management*	1,651	1,924	2,277	2,956	3,235	4,567
... *of which: Harvesting*	6,255	6,853	6,567	10,679	11,562	9,670
Total Economic Surplus$_{/\text{ha}}$ *(in KES)*	**107,002**	**117,685**	**212,284**	**80,429**	**93,303**	**129,778**

Table 6: Agronomic Value Creation across the Sample

[6] The value derived from consumption is calculated by modelling the diminishing marginal utility of potatoes eaten in the household. Consult Appendix C for details.

[7] These costs can be direct, i.e., the cost of purchasing of seed potatos, or indirect, i.e., the income lost from retaining a potato as seed, which could otherwise be sold.

4 Impact Assessment: A Roadmap

We have just shown that in our sample, the average surplus generated by an FFBS-trained farmer stands at $190,000^{KES}/_{ha}$, and $110,000^{KES}/_{ha}$ for a matched non-FFBS-trained farmer. If a member of *Treatment* is indeed comparable to a member of *Matched Control* without the influence of selection bias, it stands to reason that the per-farmer, per-season, per-hectare impact of the FFBS amounts to around $80,000 KES$. We could then simply multiply this *individual impact*, termed γ, by the number of farmers trained N, the number of seasons M the farmer will be able to apply the knowledge gained in the FFBS, and the average size of a trained farmer's potato plot A, to obtain the gross economic impact of the FFBS, termed Γ.

$$\Gamma = \gamma \times N \times M \times A \tag{1}$$

At this point, to find the net economic impact of the FFBS, we would simply subtract from Γ the total cost of the program C, and arrive at Π, the net "social surplus" generated by the FFBS over its period of activity

$$\Pi = \Gamma - C \tag{2}$$

or, alternatively, we could calculate the benefit-cost ratio π:

$$\pi = \frac{\Gamma}{C} \tag{3}$$

In many ways, this approach is correct: We will indeed aggregate γ almost exactly as in Equation 1, and calculate the net economic impact as in Equations 2 and 3. The issue is the calculation of the individual impact γ itself, which we can not simply take to be 80,000 KES, the difference between surpluses generated by *Treatment* and *Matched Control*. But why?

4.1 Impact Assessment under Imperfect Matching

The problem is that no matching model is perfect. Some selection bias always persists, even though it may well be markedly reduced by matching. For any given variable, the level of selection bias that persists in the comparison with *Matched Control* is proportional to the selection bias that exists initially, in the comparison with *Control*. Hence, to limit the amount of selection bias that persists, the researcher may focus comparison between *Treatment* and *Matched Control* on variables which are comparatively less afflicted by selection bias in the first place. How to find such variables?

The researcher argues that the more directly participation in the FFBS impacts a given variable, the lower the selection bias this specific variable is afflicted with. To understand the rationale of the argument, note that a driver x causes selection bias in the measurement of the treatment effect of the FFBS on variable y whenever two conditions arise simultaneously: *Imbalance* and *importance*. *Imbalance* arises when driver x causes heterogeneous selection into either sample group; *importance* arises when this driver x causes heterogeneous levels in variable y. Since *imbalance* is independent of the choice of variable y, variable y is afflicted by selection bias proportionally to the degree that it invites drivers which are *important*.[8]

Imagine, for example, the set of drivers that fulfill the conditions of *imbalance* and *importance* with respect to the quantity of fertilizer used during planting, and thereby afflict the estimation of a treatment effect with respect to this variable of choice, fertilization. Note that, since fertilization affects yields, any driver *important* for fertilizer quantity will also be *important* for yields. The opposite is not true: There can be drivers of selection bias in the estimation of a treatment effect of the FFBS on yields that do not affect the estimation for fertilizer quantity. An example here could be the degree of infestation with pests or diseases, which affects yields, but materializes after the administration of fertilizer, so that it can have no *importance* for it. Similarly, any driver of yields will also be a driver for revenues, but not vice versa; Finally, any driver for revenues will be a driver for surplus, but not vice versa. Therefore, the further up the chain of agronomic production one chooses a variable y, the more one invites selection bias. To combat this, we will only compare *Matched Control* and *Treatment* for those variables which are

[8] For a deeper explanation of *imbalance* and *importance*, consult Appendix B

direct effects of the FFBS. Direct effects of the FFBS include, most importantly, increases in the rate of GAP adoption. Since the GAP are directly taught as part of the FFBS curriculum, the estimation of a treatment effect with respect to GAP adoption is, by our reasoning above, least susceptible to selection bias. The set of variables encoding these direct effects is denoted $\boldsymbol{E}$. To clarify the difference between an effect and effect variable by way of example: an increase in the adoption of top dressing is a direct effect, the rate of adoption of top dressing is the associated effect variable. Any economic impact conferred by the FFBS has to be mediated by one of these effect variables.

4.2 Isolating the impact of a direct effect

What remains now is to link any given effect variable $e \in \boldsymbol{E}$ to increased farmer profits. If this link was established, we could combine it with the extent to which the FFBS causes a direct effect in e (which we find by comparing *Treatment* and *Matched Control*) to find the increase in farmer profit directly attributable to the FFBS via this effect variable e. It would then remain to sum up these increases across all effect variables in $\boldsymbol{E}$ to find the individual impact γ, from which a gross economic impact can easily be aggregated, as shown above. The missing piece, then, is the link between direct effects e and farmer profits.

To find and quantify this link, the researcher pools *Treatment* and *Control* sample, and estimates a model of agronomic value creation.[9] Details may be found in Appendix C, but the general idea is to estimate a set of equations which take in a multitude of explanatory variables $\boldsymbol{X}$ (including farmer characteristics, agricultural practices and control variables) and predict the following quantities:

1. $Y(\boldsymbol{X})$: The total yield harvested
2. $f^k(\boldsymbol{X})$: The proportion of yield that goes toward the following categories k:
   (a) sold as ware
   (b) sold as seed
   (c) retained as seed
   (d) consumed
   (e) given away
   (f) rotten
3. $V^k(\boldsymbol{X})$: The values of a kilogram of tubers belonging to each of the above categories k.

Having these expressions, the constituent parameters of which can be found in Table A9, we can calculate the revenue impact of any given direct effect variable $e \in \boldsymbol{E}$, holding all other factors constant. In the simplest case, in which our direct effect variable e is binary, the marginal revenue impact of e is as follows:

$$
\begin{aligned}
R_i^e = \sum_{k \in \boldsymbol{K}} \Big[& f^k(\boldsymbol{x} = X_i | e = 1) \times Y(\boldsymbol{x} = X_i | e = 1) \times V^k(\boldsymbol{x} = X_i | e = 1) \\
& - f^k(\boldsymbol{x} = X_i | e = 0) \times Y(\boldsymbol{x} = X_i | e = 0) \times V^k(\boldsymbol{x} = X_i | e = 0) \Big]
\end{aligned} \tag{4}
$$

Though Equation 4 may look daunting, its interpretation is relatively simple. If, for example, e is chosen to be the practice of top dressing,[10] then $R_i^{Top\ Dressing}$ expresses the revenue gained if farmer i applies Top Dressing, all else being equal. The utility of such a set of models is that we can now calculate the impact on farmer revenues directly attributable to the FFBS via (for example) Top Dressing. To obtain this figure, termed individual partial revenue impact, we simply multiply $R_i^{Top\ Dressing}$ with the increase in adoption of top dressing directly attributable

[9] Note that by pooling the two samples, selection bias due to opting in our out of the FFBS is eliminated completely in this model.

[10] which is a binary variable: either one top dresses, or one does not.

to the FFBS, termed $\Delta^{Top\ Dressing}$, which we find by comparing the rate of adoption of top dressing in the *Matched Control* with the rate of adoption in the *Treatment* sample. It then remains to subtract from the individual partial revenue impact the individual partial cost impact. The individual partial cost impact is the cost associated with applying effect variable e multiplied also by Δ^e, the rate at which the FFBS causes higher adoption of e. In the case of top dressing, this is simply the unit cost of top dressing fertilizer multiplied by a typical quantity of top dressing fertilizer applied, and then weighted by the increase in the adoption of top dressing attributable to the FFBS. Subtracting this cost impact from the revenue impact, and averaging over all FFBS-trained farmers, yields the individual partial surplus impact, or, more simply, the individual partial impact γ^e.

$$\gamma^e = \frac{1}{N_{Treatment}} \sum_{i \in Treatment} \Delta^e (R_i^e - C_i^e) \tag{5}$$

4.3 Aggregating Individual Partial Impacts

The individual partial impact γ^e is individual in the sense that it expresses an impact on a single farmer's household; it is partial in the sense that it expresses the impact of a single effect variable e. Therefore, we can aggregate it across all effect variables e to get a total gross economic impact γ, expressed in KES per farmer per season per hectare.

$$\gamma = \sum_{e \in \boldsymbol{E}} \gamma^e \tag{6}$$

Exactly as outlined in Equation 1, multiplying by the number of farmers trained, weighting by the typical size of the potato plot, and calculating a present time value across all seasons to get a lifetime impact, we finally obtain the gross economic impact of the FFBS. The gross economic impact is the total economic value created by the FFBS for the benefit of its beneficiaries; it expressed simply in KES. The only methodological difficulty in aggregation is the transition from a per-season impact to a lifetime impact. To this end, the researcher has to compound past returns and discount future returns attributable to the FFBS to express them in present time value. Only then are past and future returns commensurable, and can be added up to a lifetime impact. The methodological details of this approach are explained in Appendix D.

5 The Gross Economic Impact of the FFBS

As outlined in Section 4.2, we specify and estimate the equations necessary to determine our revenue model R_i^e. The resulting model parameters are outlined in Tables A9. In this section, we first apply our revenue model to arrive at individual partial impacts γ^e as per Equation 5. The individual partial impacts will be related, for each effect variable e separately, in Section 5.1. We will then aggregate these individual partial impacts as per Equations 6 and 1, to find the gross economic impact of the FFBS, in Section 5.2.

5.1 Individual Partial Impacts

The set of all individual partial impacts[11] is summarized in Table 7. For every relevant effect variable e, it shows the degree of adoption directly attributable to the FFBS, the *ceteris paribus* impact on total yields and average prices predicted by our agronomic models, and the resultant individual partial revenue, cost and surplus impacts. Before delving into Table 7, we will present an overview in which individual partial surplus impacts are grouped. This overview, the waterfall chart in Figure 4, takes as a baseline the seasonal profit of a member of *Matched Control*, that

[11] An individual partial impact will be outlined in the following only if the effect variable e that mediates it is revenue or cost relevant. If a direct effect does not feature in any of the revenue equations (either because we had no data on it, or because the researcher found its effect to be insignificant and negligible when settling for the set of models with the highest explanatory power, or because the variable is justifiably afflicted with bias (see section 5.1.9)), and if this direct effect e does not carry cost implications, then we can safely ignore it, as the net contribution of e will amount to zero. As the set of variables featuring in the revenue models is $\boldsymbol{X}$, relevant e fulfill $e \in \boldsymbol{E} \cap \boldsymbol{X}$. All variables $\boldsymbol{X}$ are listed in Table A9.

is, a farmer who is similar to an FFBS-trained farmer, but did not undergo FFBS training. The waterfall chart then shows how, according to our models, undergoing training would increase the farmer's profits via the application of knowledge gained in different GAP categories, or membership in a marketing group. These attributable impacts are shown in green and red. Going from left to right, in the areas of seed selection and soil fertility management, there are large surpluses that the farmer nets, surpluses that can be directly attributed to the FFBS. Land preparation and planting practices contribute slightly negatively, and no sizeable effect can be shown to accrue via Weeding, Hilling and Thinning. However, mediated by knowledge acquired about pest and disease management, and harvesting and post-harvest practices, the FFBS causes large increases in farmer surplus per season per hectare. Finally, via the promotion of marketing group membership, prices can be shown to increase, which causes another increase in surpluses due to the FFBS. However, out of of the $80,000^{KES}/_{ha}$ per season that a trained farmer takes home more than a similar untrained farmer, only around half can be attributed directly to the FFBS. The remainder is not rigorously attributable to the FFBS, and shown in grey. This non-attributable part might be caused by selection bias due to incomplete matching; it might also reflect an impact of the FFBS that is not mediated via one of the effect variables $\boldsymbol{E}$, but a different variable that we do not account for in our model. However, since the entire impact of the FFBS has to be mediated via well-specified variables that improve the predictive quality of the regression model used and thereby have an impact on yields or prices, we have to refrain attributing it, or any part of it, to the FFBS.

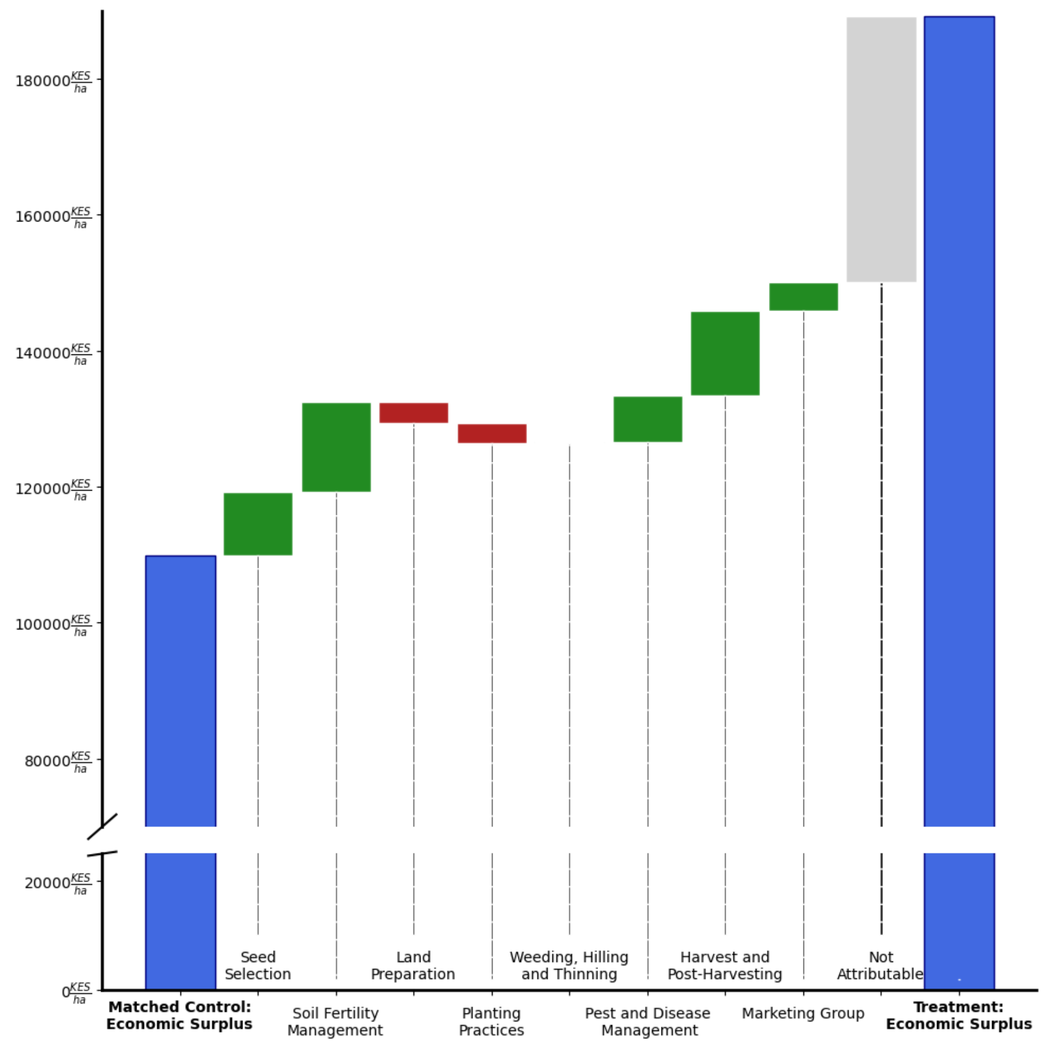

Figure 4: Individual Partial Impacts partly explain the difference between a trained and a non-trained farmer's profits.

			Adoption		Marginal Yield		Marginal Value		Marginal Revenue		Individual Partial ... Impact		
			Match. Control	Treatm.	Match. Control	Treatm.	Match. Control	Treatm.	Match. Control	Treatm.	... Revenue ... (per Season per Hectare)	... Cost ...	... Surplus ...
Seed Selection	Quality	No Selection	85%	34%									
		Pos. Selection	4%	24%	7,579	7,865	24.7	24.7	188,074	194,964	6,890	5,691	**1,199**
		Cert. or Clean	10%	41%									
	Quantity		$1311 {}^{kg}/_{ha}$	$1449 {}^{kg}/_{ha}$	7,579	8,054	24.7	24.6	188,074	199,516	11,442	4,608	**6,834**
	Interaction										3,259	1,831	**1,428**
Soil Fertility Management	Planting Fertilizer: Use		83%	93%	8,366	8,623	24.6	24.6	206,265	212,386	6,121	1,268	**4,853**
	Planting Fertilizer: Quantity		$180 {}^{kg}/_{ha}$	$216 {}^{kg}/_{ha}$	8,366	8,596	24.6	24.6	206,265	211,751	5,486	4,289	**1,197**
	Planting Fertilizer: Interaction										4,015	1,922	**2,093**
	Top Dressing: Used		32%	53%	8,847	9,025	24.6	24.6	218,550	222,801	4,251	1,625	**2,626**
	Top Dressing: Quantity		$124 {}^{kg}/_{ha}$	$162 {}^{kg}/_{ha}$	8,847	8,890	24.6	24.6	218,550	219,596	1,047	1,831	**-784**
	Top Dressing: Interaction										1,558	1,017	**541**
	Soil Testing: Recs. Applied		1%	10%	8,978	9,159	24.6	24.6	221,100	225,411	4,311	1,649	**2,662**
Land Preparation	Done before onset of Rains		80%	90%	9,329	9,203	24.6	24.6	229,861	226,852	-3,009	0	**-3,009**
	Depth		16.9 cm	16.8 cm	9,090	9,085	24.6	24.6	224,178	224,071	-106	0	**-106**
	Adequate Steps Taken		62%	67%	9,171	9,174	24.6	24.6	226,223	226,304	81	0	**81**
Planting Practices	Depth		11.2 cm	12.2 cm	9,333	9,209	24.6	24.6	230,129	227,155	-2,974	0	**-2,974**
Weeding, Hilling and Thinning	Height of Hills		22.6 cm	26.0 cm	9,172	9,159	24.6	24.6	226,265	225,968	-297	0	**-297**
	Thinning		5%	14%	9,158	9,175	24.6	24.6	225,924	226,328	403	0	**403**
Pest and Disease Management	Rogueing		17%	13%	9,198	9,188	24.6	24.6	226,900	226,651	-249	0	**-249**
	Disease Chemicals		56%	76%	8,871	9,146	24.6	24.6	219,305	225,872	6,567	1,439	**5,128**
	Pest Chemicals		21%	37%	9,101	9,193	24.6	24.6	224,736	226,920	2,184	127	**2,057**
Harvest and Post-Harvesting	Dehaulming		15%	64%	8,691	9,152	24.6	24.6	213,935	224,890	10,956	0	**10,956**
	Sorting and Grading		87%	96%	9,175	9,175	24.4	24.6	224,586	226,114	1,528	0	**1,528**
Marketing Group	Membership		2%	24%	9,175	9,175	24.1	24.6	221,620	225,878	4,258	0	**4,258**
											67,722	**27,297**	**40,425**

Table 7: Drivers of Gross Economic Impact

We will now go through Table 7, effect variable by effect variable, and discuss the atomic constituents of each green or red bar in the waterfall chart in Figure 4.

5.1.1 Seed Selection

The impact mediated via effect variables related to seed is driven in good part by the adoption of positively selected and certified or clean seed, where increases of 20 and 30 percentage points respectively are observed between *Matched Control* and *Treatment*. Keeping seed quantity constant, these FFBS-attributable increases in seed quality contribute $6,890 {}^{KES}/_{ha}$ in increased farmer returns, every season. However, the strongest driver of increased revenues are increases in quantity: A trained farmer plants $138 {}^{kg}/_{ha}$ more than his matched untrained counterpart. This increase in seed rate alone, at constant seed quality, causes yields to increase by almost $500 {}^{kg}/_{ha}$, which translates to an increase in revenue of $11,442 {}^{KES}/_{ha}$. Additionally, there is an interaction effect between the FFBS impact on quality and quantity,[12] which further contributes $3,259 {}^{KES}/_{ha}$ in increased farmer revenues attributable to the FFBS. To obtain this higher quality seed, a farmer incurs costs of $5,691 {}^{KES}/_{ha}$, which offset a good part of the gains from higher quality seed. The cost of higher seed rates is slightly lower, at $4,608 {}^{KES}/_{ha}$, much lower than the revenue conferred by increasing seed rates. Lastly, the revenue from the interaction effect is contrasted with an interaction cost, to find a total surplus per hectare per season attributable to the FFBS via seed selection, which amounts to $1,199 + 6,834 + 1,428 \approx 9,460 {}^{KES}/_{ha}$.

For a deeper look at the gross economic benefit created by the FFBS via seed selection, consider Figure 5. Each of the panels corresponds to a degree of seed quality: from left to right, we have non-selected seed, positively selected seed, and clean or certified seed. For each of these qualities, we compare the marginal revenue with the marginal cost, shown in green and red respectively. The marginal revenue tells us how much revenue is gained by increasing the seed rate by one ${}^{kg}/_{ha}$. The marginal cost tells us how much additional cost is incurred by increasing the seed rate by one ${}^{kg}/_{ha}$. The marginal cost is constant: Each additional kilogram costs a certain unit price. The marginal revenue is non-linear, initially very large and then tending towards zero. If marginal revenue is above marginal cost, each additional kilogram generates a profit. The graph shows well why a large part of the impact caused by seed selection is due to increases in quantity. For each of the three seed quality steps, the seed rate of the *Matched Control* group (marked light grey) is lower than the seed rate of the *Treatment* group (dark grey). Since, at "normal" seed rates, marginal revenue is well above the marginal cost, higher seed rates translate into increased profits.[13]

[12] i.e., farmers that use higher quality seed simultaneously use higher seed rates.

[13] In fact, the value of the profit gained can be calculated by integrating marginal revenues and costs between

Economic logic dictates that a farmer keep increasing his seed rate as long as marginal revenue is above marginal cost: This way, each additional kilogram of seed generates a profit. The optimal quantity of seed is reached where marginal revenue equals marginal cost. This point is not reached within Figure 5 for any seed quality: Our model values higher seed rates very highly.

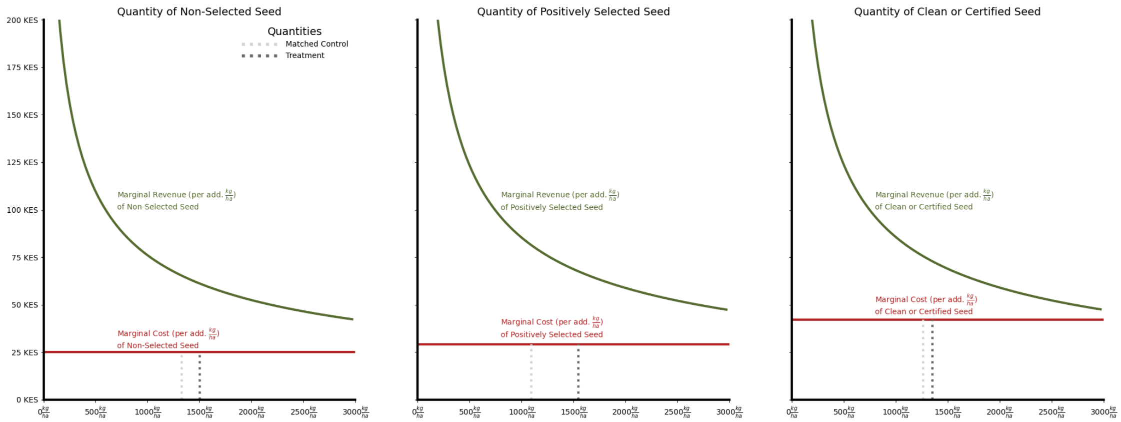

Figure 5: Marginal Revenues and Costs of Seed

5.1.2 Soil fertility management

Mediated by learning in soil fertility management, the FFBS causes a seasonal impact on farmer profits of around $13,000^{KES}/_{ha}$. Most of this is driven by the use of fertilizer at planting. Both the higher adoption of fertilizer use (up by 10 percentage points between *Matched Control* and *Treatment*) and the quantity of fertilizer used contribute to this figure, with a larger impact on the part of the former. Trained farmers use an additional $36^{kg}/_{ha}$ of fertilizer, which causes, all things equal, an increase in yields by $230^{kg}/_{ha}$. Additionally, a sizeable interaction effect captures the joint impact of higher adoption and increases in quantity. In total, the revenue increase attributable to the FFBS via fertilization at planting stands at more than $15,000^{KES}/_{ha}$, and well exceeds the associated cost increase of around $7,500^{KES}/_{ha}$.

The impact of top dressing is more subdued, and comes entirely from the adoption channel: 53% of trained farmers top dress, as opposed to 32% of farmers in the *Matched Control* group. The increase in quantity of top dressing fertilizer used is sizeable, at $38^{kg}/_{ha}$; however, revenue gained from this increase in isolation is subdued, at around $1,000^{KES}/_{ha}$, and the surplus is negative. Figure 6 shows why. On the right hand side panel, we compare the marginal revenue with the unit cost of a kilogram of top dressing fertilizer. The optimal rate of top dressing predicted by the model stands at circa $100^{kg}/_{ha}$. *Matched Control* group members are very close to this rate, *Treatment* group members somewhat overshoot it. Even though the revenue gained is still positive for a *Treatment* group member, it is rapidly decreasing, which explains the subdued additional revenue from top dressing quantities. In fact, any unit of fertilizer beyond the predicted optimum actually costs more than it earns, so that, after factoring in costs, the quantity impact of top dressing fertilizer is, on the whole, negative. Note, however, that together with the effect of a higher adoption rate, the overall impact attributable to the FFBS via top dressing is still positive.

On the other hand, the left hand side graph shows that the impact attributable to the FFBS via the quantity of planting fertilizer is net positive. The optimal quantity of planting fertilizer is predicted by the model to lie at around $320^{kg}/_{ha}$, above the rate used by both *Matched Control* and *Treatment*. However, this time, *Treatment* quantities are closer to the optimum, so that the quantity channel carries a net positive impact.

the two seed rates, and subtracting these two integrals. Essentially, the increase in profit corresponds to the size of the area enclosed by the two curves and the levels of adoption.

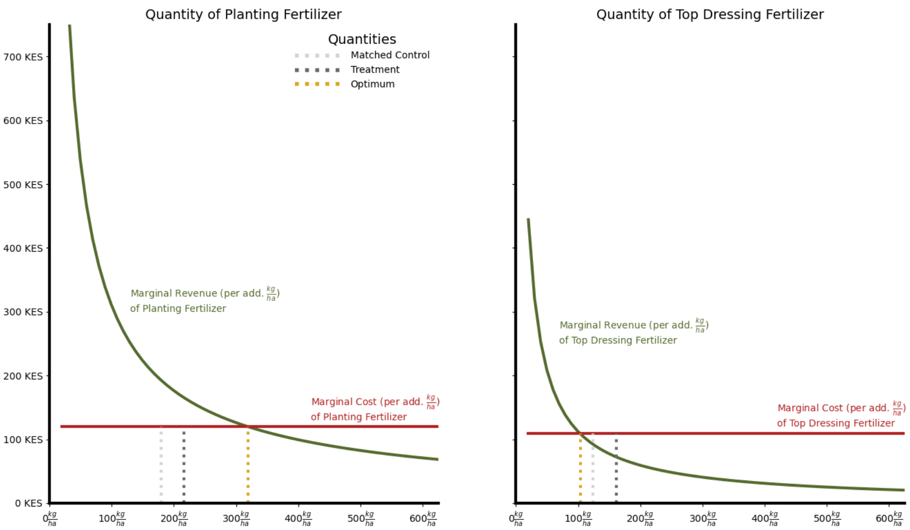

Figure 6: Marginal Revenues and Costs of Fertilizer

An additional driver of the FFBS impact mediated by soil fertility management is the application of soil testing recommendations. The yield model includes a binary variable that equals one if soil testing is done and the recommendations are fully applied; and zero otherwise. According to the model, fully applying the recommendations of the soil test increases yields by slightly over 18%.[14] Combined with a modest increase in adoption - in the *Treatment*, 10% soil tested and applied the recommendations fully, up from virtually no one in *Matched Control* - the partial revenue created by the FFBS via soil testing amounts to $4,311^{KES}/_{ha}$, netted against a cost of $1,649^{KES}/_{ha}$ to yield a surplus impact of $2,662^{KES}/_{ha}$.

Note that the cost figure is calculated for *all* people taking the test, not just those fully applying its recommendations. Even after accounting for the attrition occurring due to people taking the test, but not implementing its recommendations fully, the net surplus created by the FFBS here is positive. As can be seen in Table 8, only about 25% of test takers fully apply the recommendations, so that the marginal cost reported in Table 7 is four times as large as it would be at theoretical full compliance. In practice, this attrition is countered by the fact that some soil tests are gifted - among those that took the test, only 44% had paid for it. However, to be conservative, we calculate the cost of soil testing only on the basis of those that actually paid for the test.

Interestingly, receiving a free test reduces the propensity to apply the recommendations drastically, as can be seen in Table 8. This might be due to a variety of factors. Farmers that pay for the test may be ex ante more eager to implement its recommendations; conversely, some farmers receiving a free test may not intend to follow up on it in the first place. Additionally, paying (even a subsidized fee) for a test may nudge a farmer toward applying its recommendations due to a sunk-cost effect (Arkes & Blumer, 1985); Finally, higher cost may be seen as indicative of the test's quality, and thereby increase propensity of application (Riley, 2001). On the other hand, the reduction in testing due to higher prices may partially or fully negate the effects of higher rates of application. Consult Cohen and Dupas (2010) for a good review of perspectives and results on the matter, along with an empirical test of difference between free and subsidized provision of mosquito nets in Kenya.

[14] Due to the need to linearise multiplicative models to prepare them for estimation, an estimated parameter of 0.1865 (see Table A9) actually implies an increase in yields by $e^{0.1865} - 1 \approx 20.5\%$. However, as is the case for all estimated parameters not too different from 0, the direct approximation of 18.65% requires no computation and is close to the real estimated model value.

	Did not apply	Partly applied	Fully applied
Soil test was gifted	46%	36%	19%
Soil test was paid	14%	54%	32%
All	**32%**	**44%**	**25%**

Table 8: Those who paid for a soil test are more likely to apply its recommendations

5.1.3 Land Preparation

GAP 7 specifies that the land be prepared before the onset of rains. The estimated *Treatment* effect for adoption of this GAP is estimated to stand at ten percentage points. However, including this variable in the model results in a negative coefficient for yield that is robust against multiple different specifications. Therefore, this increase in adoption causes a drop in predicted yields, and therefore revenue. The impact is not large, amounting to around $3,000^{KES}/_{ha}$. The revenue impacts of GAP 8 and 9, i.e. the depth of land preparation and the correct sequence of steps taken,[15] are negligible. For the adequate steps taken in land preparation, this is due to the model estimating a minuscule impact on yields gained from GAP adherence. On the other hand, the average depth of land preparation is virtually identical in *Matched Control* and *Treatment*, so that the revenue impact is very small. However, the model does estimate a large preference for deeper land preparation, as shown in the left hand panel of Figure 7: For the entire range of preparation depths observed, the marginal benefit is resoundingly positive. Do note that the marginal cost for an additional centimeter of tillage is very likely small, but positive. However, the researcher could not establish it by regressing land preparation labor costs on land preparation practices, including the depth of tillage. Hence, the marginal cost of an additional centimeter of tillage is assumed to be zero, as in the other panels in Figure 7.

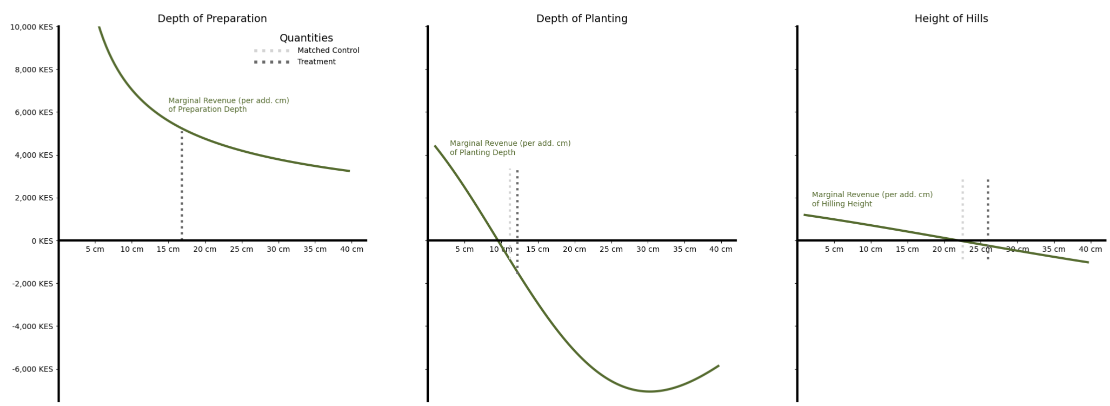

Figure 7: Marginal Revenues of Land Preparation and Planting Depth, and Hilling Height

5.1.4 Planting Practices

Figure 7 also explains why the revenue impact attributable to the FFBS is moderately negative for the depth of planting, and insubstantially negative for the height of hills. *Matched Control* members plant their tubers at on average 11.2 cm, and *Treatment* members at 1 cm deeper. However, the model predicts an optimal planting depth at around 10 cm, where marginal revenue crosses the x-axis; any additional depth contributes negatively to revenue, causing a negative revenue impact attributable to the FFBS.

[15] For old, fallow or virgin land, at least two, three or four (respectively) of the following four steps have to be taken: 1) clearing (mechanically or with herbicide), 2) primary ploughing or chiseling, 3) secondary ploughing, 4) furrowing or harrowing

5.1.5 Weeding, Hilling and Thinning

Similarly, the optimal height of hills is predicted to be around 22.5 cm, almost exactly the height at which the average *Matched Control* member hills up. The average *Treatment* group member, on the other hand, hills up higher, at 26 cm, hence the revenue impact is again negative. However, since the marginal revenue curve is very flat, the revenue impact is small.

Thinning is predicted by the model to increase yields by a very modest 2%. As the adoption of the practice between *Matched Control* and *Treatment* increases by around 9%, Thinning contributes around $400^{KES}/_{ha}$ in increased seasonal farmer revenues attributable to the FFBS.

5.1.6 Pest and Disease Management

As with thinning, rogueing is predicted to modestly increase yields by the model, by around 3%. However, the adoption of rogueing slightly decreases, causing a negative, but small, revenue impact. On the other hand, the increased application of disease and pest control chemicals attributable to the FFBS causes a large revenue impact. The prevalence of using chemicals to combat pest and disease increases by 20% and 16% respectively. It then remains to combine these matched adoption figures with the increase in yields derived from chemical pest and disease management. Our yield model estimates just this increase, while controlling for the severity of the infestation - if this was not done, we would likely find a negative yield effect,[16] since the application of pesticide correlates with the degree of infestation, which in turn causes lower yields. Via chemical disease management, we find the FFBS to contribute more than $250^{kg}/_{ha}$ in increased yields and around $6,500^{KES}/_{ha}$ in increased revenue; via chemical pest management, yields are boosted by almost $100^{kg}/_{ha}$ and revenues by almost $2,200^{KES}/_{ha}$ by the FFBS, with the increases in revenue dwarfing the corresponding increases in costs due to the procurement of chemicals and the labor cost of application.

5.1.7 Harvesting Practices and Post-Harvest Management

Harvest and post-harvest management is an area in which the knowledge imparted by the FFBS causes substantial surpluses enjoyed by the beneficiaries. As we see in Table 7, these benefits are mediated mainly by the practice of dehaulming, which, between *Matched Control* and *Treatment*, sees a staggering 50 percentage point increase in adoption. This increase, combined with a yield model which estimates that the practice of dehaulming increases harvested yields by 10% *ceteris paribus*, causes an increase in revenue attributable to the FFBS of almost $11,000^{KES}/_{ha}$ per farmer, per season. An additional driver of impact within harvest and post-harvest management is the practice of sorting and grading. Though already wide-spread in the *Matched Control* group, adoption increases to virtually everyone in the *Treatment* sample. The price model predicts that the ware price fetched for sorted and graded potatoes increases by 7%, all else being equal. Additionally, sorted and graded potatoes are more likely to be kept or sold as seed as opposed to ware, further increasing the harvest's average value. Hence, even a single-digit adoption increase contributes around $1,500^{KES}/_{ha}$ in revenue impact attributable to the FFBS.

5.1.8 Marketing Group Membership

Our model also takes into account marketing group membership, which is predicted to boost prices of tubers sold by more than 8%. Since virtually no *Matched Control* respondent is member of a marketing group, versus almost a quarter of respondents in the *Treatment*, the revenue impact is sizeable, at more than $4,000^{KES}/_{ha}$.

5.1.9 Notable Omissions

The most striking omission among all candidate effect variables is the effect of crop rotation. The model does indeed include a related variable: A binary regressor indicating whether potato was not planted on the main plot (on which, in the current season, potato was planted) in the season before the current season. In the yield model regression, one would expect the coefficient of this regressor to be positive: Not planting potato twice (or more) in a row allows nutrients to

[16] as occurred in the previous economic impact assessment (Vagliano, 2019).

replenish. However, the regressor is negative and very large: The model seems to suggest that, all else being equal, a farmer who previously had planted potatoes on the field in question will have a yield higher by around 26.9%, compared to a farmer who had something else planted previously. The problem here is that planting potatoes twice in a row correlates with the natural fertility and fitness of the soil for potato farming. Consider a farmer who plants potato in the season before the current and sees staggering yields: she is much more likely to plant potatoes on the same plot in the next season, owing to the tremendous past harvest. In the next season, the season for which we have collected data, her yields will likely again be very large, though possibly slightly reduced compared to last season, due to some soil depletion. However, we only observe this season's high yields. Hence, what we are actually measuring is not the (detrimental) effect of planting potato twice in a row, but the current fertility of the plot of land with respect to potatoes. The researcher did try to control for it, by including as a control variable in the yield equation the yields per hectare of farmers in the immediate vicinity of the plot in question; this serves to reduce the magnitude of the negative coefficient, but does not completely eliminate the omitted variable bias, that is, the confounding effect introduced by not being able to control for a variable (current fertility) correlated both with the independent variable (yields per hectare) and the regressor in question (not planting potato twice on the same plot in subsequent seasons). In the light of this, the binary crop rotation regressor is left in, but serves as a proxy for fertility, controlling for what we can not observe.

In part due to this issue, the effect of nutrition training on yields is very hard to estimate. One can link nutrition training to the propensity to rotate crops: If nutritional training encourages food diversity, then farmers trained in the nutrition-integrated FFBS are more likely to rotate crops, since it promotes a more varied diet. However, because of the conundrum above, no effect on yields can be established via this vector. Additionally, including the individual dietary diversity score (IDDS) as an explanatory variable in the yield model does not result in a significant parameter. For a deeper dive into the possible vectors of economic impact that the nutrition component of the integrated FFBS may have, consult Appendix F.

To some extent, the benefits of social cohesion are captured by the direct effect of marketing group membership; however, these groups are by nature concerned with potato promotion and sale, and not agricultural practices. Therefore, we are additionally interested in the benefit conferred by remaining active within one's FFBS group after training activities have concluded. After all, remaining active here might aid in knowledge retention of the GAP, or alternatively cause persistently elevated GAP adoption via implicit or explicit peer pressure. This hypothesis can be tested, but not as a stand-alone partial impact. Under the hypothesis, the effect of group cohesion is to maintain an elevated level of GAP adoption over time, thereby potentially amplifying the effects of all other individual partial impacts in seasons beyond the end of training. We model this reduction in adoption decay directly, by attempting to estimate a decay parameter p' for those farmers still active within their groups and those not meeting their groups anymore separately. The results of this estimation, and the way it allows us to calculate the economic value of group cohesion, are outlined in Section 5.2.2. The specification and estimation of the decay model is given in Appendix E.

5.2 Gross Economic Impact

The gross economic impact of the FFBS is the total economic value created by the FFBS for the benefit of its beneficiaries. As outlined in Section 4.3, having calculated the individual partial impacts, we can easily aggregate to find the total value. This calculation is outlined in Table 9, and commented upon briefly in the following section.

5.2.1 Aggregating Individual Partial Impacts

To calculate the gross economic impact, we first sum up the individual partial impacts, that is, the surpluses attributable to the FFBS via all relevant effect variables. This yields the individual, seasonal impact figure of $40,425^{KES}/_{ha}$ seen already in Table 7. We then multiply this figure by the area under potato for each farmer (on average, 0.29 ha) and compute the mean, to obtain an

Individual Partial Impacts	
Seed Selection	9,460 KES
+ *Soil Fertility Management*	13,188 KES
+ *Land Preparation*	-3,034 KES
+ *Planting Practices*	-2,974 KES
+ *Weeding, Hilling and Thinning*	106 KES
+ *Pest and Disease Management*	6,937 KES
+ *Harvest and Post-Harvesting*	12,484 KES
+ *Marketing Group*	4,258 KES
Gross Economic Impact (per Farmer, per Season, per Hectare)	**40,425 KES**
× *Average size of potato plot*	0.29 ha
Gross Economic Impact (per Farmer, per Season)	**10,875 KES**
× $M(p, r)$, *where*	9.7
... p	15%
... r	8%
Gross Economic Impact (per Farmer)	**105,488 KES**
× *FFBS Beneficiaries*	17,927
... *of which: in Nyandarua*	11,908
... *of which: in Bungoma*	2,901
... *of which: in Trans Nzoia*	2,312
... *of which: in Elgeyo Marakwet*	806
Gross Economic Impact	**1,891,074,413 KES**

Table 9: Aggregating the Gross Economic Impact of the FFBS

individual, seasonal impact of 10,875 KES.[17] Next, we discount the seasonal impact to obtain a lifetime impact. As derived in Appendix D.1, we apply the present value multiplier of $M = 9.7$ to arrive at a gross economic impact per farmer of 105,488 KES. Multiplying this individual figure with the overall number of farmers trained (17,927), we obtain a gross economic impact of 1,891,074,413 KES or 15,726,060 EUR.[18]

5.2.2 A Perspective on Social Sustainability: The Value of Group Cohesion

Beyond the effect of marketing group membership, which is modelled as a direct effect, the researcher is interested in the extent to which remaining in the FFBS even after training activities have been concluded creates economic value. As outlined in Section 5.1.9, the impact of FFBS group cohesion can not readily be expressed as an individual partial impact. Instead, it is more elegantly modelled as a mechanism that prevents the decay of GAP knowledge and adoption: After all, one could argue that continuing to meet within the FFBS group even after the end of training helps to keep the knowledge gained in the farmer's mind. The group will likely continue discussing the GAP, and thereby reinforce individual knowledge; additionally, there might be a degree of implicit peer pressure due to in-group comparison of agricultural practices. Indeed, there seems at least to be a preference of trained farmers to remain active within one's FFBS after training has concluded: As shown in Figure 8, group persistence in the FFBS is high, at slightly more than 80% in *Treatment*, on average. This means that four out of five beneficiaries are still meeting with their FFBS group, be it pure or nutrition-integrated. Additionally, around 5% of beneficiaries meet other, non-FFBS farmer groups, for a total of around 85% rate of group membership in the *Treatment*. Compared to the *Matched Control*, where only 8% of farmers meet in any farmer group, this is a staggering, tenfold improvement in the rate of group membership, mostly facilitated by the cohesion of FFBS groups.

However, does this considerable degree of group cohesion create tangible value by preventing

[17] Note that this figure is not exactly equal to $40,425 ^{KES}/_{ha} \times 0.29$ ha, since the individual partial impacts correlate with farm size. In this case, taking the mean is a non-linear operation, and the average of products may be different from the product of averages, by Jensen's inequality.

[18] Here, and subsequently, the exchange rate used is as of the 31.08.2022 (EUR-KES: 120.25), the date corresponding to the "present" for discounting purposes; see Appendix D.1 for details.

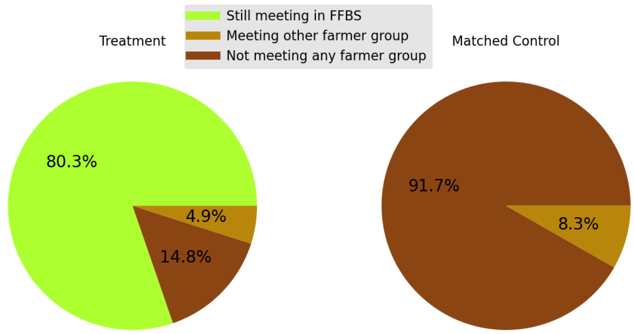

Figure 8: Farmer group cohesion and membership

the decay of GAP knowledge and adoption? In Appendix E, we estimate the degree by which GAP knowledge and adoption decays, season-by-season, after training has ended. We do this separately, for those still meeting within their FFBS group, and those not meeting. For those still active in their group, we find no adoption decay significantly different from zero; for those who are not meeting in their group, around 6.6% of GAP knowledge gained during the FFBS training is lost, each season.

As can be seen in Table 9, a very important driver of gross economic impact is the multiplier M, which translates seasonal to lifetime economic impact. This multiplier depends on parameter p, which expresses the probability that the benefits conferred by the FFBS via training are not realized anymore, beginning from the next season. We can now use our base parameter p, developed in Appendix D.1, and augment it by the increased rate of adoption decay of inactive former FFBS members. This allows us to calculate two separate multipliers M: For those active within their groups, the multiplier M is as above: 9.7. For those not meeting within their groups, the multiplier is lower, at 8.6.[19] Plainly, this means that the benefits of the FFBS are conferred to active members of their group for slightly more than a full season longer, measured at present time value. If no farmer was still meeting in his or her group, the gross economic impact reported above would be reduced by slightly more than 10%, or around 210,000,000 KES, which can be read as the value of group cohesion.

6 The Cost of the FFBS

To arrive at a net economic impact, the cost of the FFBS has to be contrasted with the gross economic impact figure developed above. However, defining a single cost figure is not sufficient: Different stakeholders have different notions of cost that they regard relevant. For a party looking to replicate the program, only the cost of actually implementing the FFBS is of importance; the expenditures incurred in developing the concept, curriculum and materials are sunk, and do not have to be borne again. In line with this view, a *Partner-Perspective* cost will be outlined first, which focuses only on the costs of implementation of the FFBS as carried out in Kenya between 2016 and 2022. This perspective is especially relevant for the continuation of the FFBS methodology, through e.g. county officials. On the other hand, the party that bore the cost of development, that is, GIZ, is interested in this sunk cost. To this end, a *GIZ-Perspective* cost will be outlined, which includes all expenditures actually booked within the accounting system of the General Project (GP) directly or indirectly attributable to the FFBS. Finally, to account for a societal, ex-ante perspective, a *Total Development and Implementation* cost figure will be calculated, which includes every cost item regardless of the interests of a particular stakeholder.

[19] The keen reader will wonder why the multiplier for trained farmers still meeting within their groups is the same as the multiplier for all trained farmers. As shown in Appendix E, the decay parameter p' can not be shown to be different from zero for both the trained farmers still meeting and all trained farmers. Hence, for both of these samples, $p' = 0$ and $M = 9.7$.

6.1 Partner-Perspective Costs

The bulk of *Partner-Perspective* costs consists of expenditures directly attributable to the running of a single FFBS group for the duration of the standard 15 sessions. These costs, referred to in the following section as variable costs, are listed in Table 10.

The variable cost figure is calculated separately for an edition of a staff-led FFBS, and a farmer-led FFBS. In a staff-led FFBS, a ward agricultural officer (WAO) teaches the FFBS curriculum, whereas a farmer-led FFBS is held by a trained farmer. The former receives an allowance of 2,040 KES, and costs an additional 3,927 KES in daily gross wages, social security contribution, and so on; the latter receives an allowance of 800 KES, for each of the 15 sessions. To ensure a high standard of training, each farmer-led FFBS is backstopped by a WAO for 5 sessions, at the WAO rates and wages outlined above. The cost of agricultural inputs differs between the staff-led and the farmer-led FFBS mostly due to the size of the demo plot: All staff-led FFBS are held on a 0.25 acre ($\approx$ 0.1 hectare) plot, whereas the farmer-led FFBS take place on a plot half that size. Conversely, the cost of stationery (which includes training materials and supplementary items) differs due to the difference in number of participants between a staff-led and a farmer-led FFBS: On average, 14 people attended a farmer-led FFBS, versus 24 in a staff-led FFBS. All in all, the gross variable cost incurred in running a 15-session group amounts to almost 120,000 KES for a staff-led FFBS, and 60,000 for a farmer-led group.

The gross variable cost does not take into account items for which no explicit expense occurs. Nevertheless, these cost items can be economically relevant, and are therefore included in a net figure. For one, the demo plot is offered by a member of the FFBS, at no charge. A calculatory item covers the cost of a comparable plot of land, at 22,500 $^{KES}/_{ha}$ rent per season. Similarly, the time and labor of participants is valued at 300 $^{KES}/_{day}$, over a total of 15 sessions. On the other hand, the value of tubers harvested, which are distributed to the participants for use as high quality seed potatoes, offsets some of the cost incurred in running the FFBS. Going by demo plot harvest data, which yield almost 21 $^{tons}/_{ha}$ on average, the harvested tubers, valued at a conservative 25 $^{KES}/_{kg}$, compensate some, but not all of the participant's time and labor invested. All in all, the net variable cost of running a single staff-led FFBS amounts to around 125,000 KES, and just under 100,000 KES for a farmer-led FFBS.

To the total variable costs, which are scaled up proportionally to the number of FFBS held in each year, are added costs related to project coordination and the training of trainers. These are not directly attributable to the running of a single FFBS, and therefore calculated season-by-season, as shown in Table 11.

	Variable Costs			Coordination Costs		Training of Trainers Costs				**Partner-Perspective Costs**	
Year of Training	*No. of Staff-Led FFBS*	*No. of Farmer-Led FFBS*	Total	County and Subcounty Team	Project Management	*WAOs Trained*	Cost of Training WAOs	*Farmers Trained*	Cost of Training Farmers	Total	**Total at Present Value**
2016 B	*0*	*0*	0	0	0	*23*	2,471,961	*0*	0	2,471,961	**5,763,720**
2017 A	*19*	*0*	3,333,259	713,782	482,891	*21*	1,877,182	*57*	713,464	7,120,577	**15,372,791**
2017 B	*18*	*57*	8,686,000	1,070,673	536,367	*27*	2,347,491	*52*	962,518	13,603,050	**27,192,560**
2018 A	*62*	*54*	16,114,169	1,070,673	1,324,047	*0*	0	*132*	2,757,411	21,266,300	**39,362,437**
2018 B	*18*	*132*	15,959,916	1,070,673	536,367	*20*	1,132,782	*60*	959,109	19,658,847	**33,691,810**
2019 A	*55*	*62*	15,662,011	1,665,491	1,317,698	*20*	1,132,964	*51*	1,044,855	20,823,018	**33,043,513**
2019 B	*44*	*116*	18,969,449	1,665,491	1,120,778	*0*	0	*0*	0	21,755,718	**31,966,287**
2020 A	*0*	*0*	0	0	0	*0*	0	*0*	0	0	**0**
2020 B	*76*	*0*	13,333,035	1,665,491	1,693,636	*0*	0	*50*	682,023	17,374,185	**21,886,469**
2021 A	*0*	*91*	8,825,685	1,665,491	333,098	*0*	0	*0*	0	10,824,274	**12,625,433**
2021 B	*51*	*0*	8,947,168	1,776,524	0	*0*	0	*0*	0	10,723,692	**11,581,587**
2022 A	*65*	*71*	18,289,227	2,109,622	0	*40*	1,902,891	*0*	0	22,301,740	**22,301,740**
											254,788,348

Table 11: Aggregating Partner-Perspective Costs (in KES, except where *italic*) across Seasons

With the exception of Season 2020A, in which all activities were halted due to the Covid-19 pandemic, 15 coordination meetings are held by each county and subcounty coordinator (collectively termed CC) each season in each county actively implementing the FFBS. The sum total of their allowances and daily wage costs (including tax, social security contributions, etc.) constitutes the costs booked under County and Subcounty Team coordination. Additionally, three meetings per season per county are held with all WAOs and CCs in attendance; the sum total of allowances and wage costs constitutes the costs booked under Project Management. The remaining cost item is concerned with all expenditures incurred in the training of trainers. At irregular intervals, WAOs and lead farmers have to be trained and re-trained to hold staff-led

	Staff-Led	Farmer-Led
Labour	**89,509 KES**	**41,836 KES**
Trainer	89,509 KES	12,000 KES
Daily Wage Cost	*3,927 KES*	
Training per Diem	*1,540 KES*	*500 KES*
Transport per Diem	*500 KES*	*300 KES*
Sessions	*15*	*15*
Backstopper		29,836 KES
Daily Wage Cost		*3,927 KES*
Training per Diem		*1,540 KES*
Transport per Diem		*500 KES*
Sessions		*5*
Agricultural Inputs	**20,000 KES**	**11,200 KES**
Seed	12,000 KES	6,000 KES
Seed Quantity	*200 kg*	*100 kg*
Seed Cost	*60 KES/kg*	*60 KES/kg*
Fertilizer	3,000 KES	1,700 KES
Soil Test	1,000 KES	1,000 KES
Pesticide	3,000 KES	1,500 KES
Distribution Costs	1,000 KES	1,000 KES
Stationery	**8,375 KES**	**6,025 KES**
Note Books	1,250 KES	750 KES
Pens	625 KES	375 KES
Markers	100 KES	100 KES
Rulers	100 KES	100 KES
Flip Charts	600 KES	600 KES
Farmer's Manual	4,000 KES	2,400 KES
Trainer's Manual	1,700 KES	1,700 KES
GROSS VARIABLE COST	**117,884 KES**	**59,061 KES**
Land	**2,278 KES**	**1,139 KES**
Area	*0.1 ha*	*0.05 ha*
Cost per Season	*22,500 KES/ha*	*22,500 KES/ha*
Labor and Time of Participants	**107,702 KES**	**63,000 KES**
Calculatory per diem	*300 KES*	*300 KES*
Average Number of Participants	*24*	*14*
Sessions	*15*	*15*
Value of Goods Produced	**52,430 KES**	**26,215 KES**
Yield Harvested	$20,972^{kg}/_{ha}$	$20,972^{kg}/_{ha}$
Area	*0.1 ha*	*0.05 ha*
Value of Tubers	$25^{KES}/_{kg}$	$25^{KES}/_{kg}$
NET VARIABLE COST	**175,435 KES**	**96,986 KES**

Table 10: The Variable Cost of running one FFBS for one Season

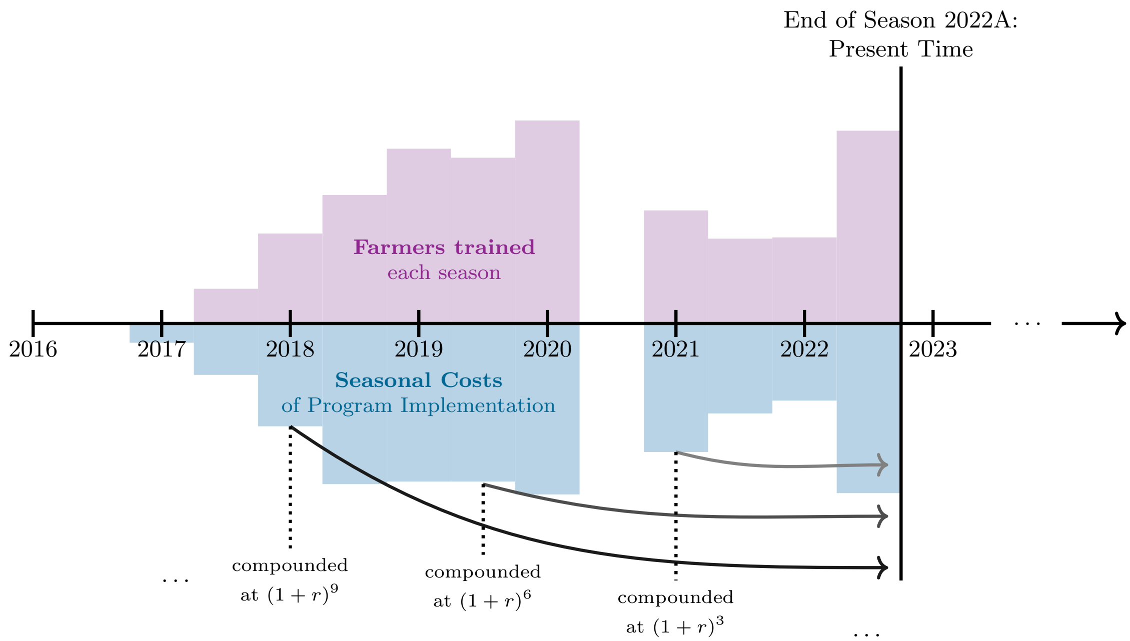

Figure 9: Calculating the present time value of program costs

and farmer-led FFBS trainings respectively. These events take place at large venues with up to 100 prospective trainers in attendance, trained over a span of up to two weeks by typically two master trainers. Venue and catering costs are considerable. Master trainers receive a salary, plus night and day travel allowances; additionally, a transport flat rate is paid to master trainers and participants alike. All these expenses contribute to the costs of training trainers reported in Table 11.

All cost items sum up to a total seasonal cost, which is stated in the respective season's nominal currency terms. Since money itself changes value over time, seasonal costs have to be expressed in terms of the value of currency at a fixed time point. Within the scope of this study, for benefits and costs alike, this time point is chosen to be the end of August 2022, which marks the end of season 2022A, which our survey pertains to. Compounding up to this time, at the rate of $r = 8\%$ seasonally,[20] yields the final column in Table 11, the total seasonal cost measured at present value. A schematic representation of this compounding schedule is displayed in Figure 9. Since seasonal costs are now expressed in like terms, they can be validly summed up, to yield a total partner-perspective cost of $254,788,348$ KES ($2,118,804$ EUR), or $14,213$ KES (118 EUR) per beneficiary.

6.2 GIZ-Perspective Costs

As opposed to the partner, GIZ is interested in costs beyond the pure implementation of the FFBS. The costs of developing the curriculum, monitoring its implementation, and analysing its results ought to be included from the perspective of GIZ. To this end, all expenditure data[21] from project accounting were collected for the years between program inception and the present. However, these expenditures were undertaken for the entire Kenya county component of the GP, of which the FFBS is only one field of action. Hence, some disaggregation is necessary.

Recall from the introduction that the GP is implemented across three fields of action, concerned with 1) improving the quality and quantity of potato production and marketing, 2) improving nutritional habits and basic hygiene, and 3) improving coordination within the potato sector. For each of these fields of action, a desired outcome is provided. To facilitate the success of these outcomes, four deliverable outputs are defined. These are concerned with a) small-scale

[20] See Section D.1 for a discussion on this rate

[21] Net of 19% VAT.

farmer GAP application, b) access to quality seed potatoes, c) nutritional knowledge especially in mothers and mothers-to-be, and d) public-private strategic planning at the national level.

Note that the FFBS program is almost perfectly congruent with Output a): On the one hand, all expenditures undertaken to meet Output a) serve to fund the FFBS. The obverse is not quite true: Expenditures toward the nutrition components of the nutrition-integrated FFBS are booked under Output c). Hence, by focusing on expenditures accrued under Output a), we find the cost associated with the agricultural component of all 991 FFBS held between 2017 and 2022. We focus on the agricultural component because the economic benefit of the nutritional components of the integrated FFBS could not be ascertained. Hence, the gross economic impact of the "pure" FFBS will be compared to the cost of its development and implementation.

Where expenditures could be clearly attributed by a certain percentage to Output a) of the GP, expenditures were counted accordingly; where expenditures were shared among the four outputs, a quarter of the expenditure was counted towards the FFBS. This data was then aggregated to arrive at a single cost figure attributable to the FFBS: $1,552,859$ EUR. To this, 14.5% were added to account for GIZ overhead, to arrive at $1,778,024$ EUR, or $213,809,124$ KES. Since expenditures could not be mapped to single seasons, costs were assumed to be distributed in line with *Partner-Perspective* costs: If, for instance, in a given season, the amount spent on implementing the FFBS amounts to 10% of the nominal total cost, then the *GIZ-Perspective* cost in that season is assumed to amount to 10% of $213,809,124$ KES. Seasonal costs so distributed are then compounded up to present value, and summed up, as is done in the leftmost set of columns in Table 12, to arrive at $324,410,332$ KES ($2,697,777$ EUR), or $18,096$ KES (150 EUR) per beneficiary.

6.3 Total Development and Implementation Costs

Table 12 also collects all cost items that are missing in the *GIZ-Perspective*, but accounted for within the *Partner-Perspective*, to find a *Total Development and Implementation* cost. This cost perspective contains every expenditure and non-accounting item, be it found in either of the two perspectives, or in both, thus reflecting all development and implementation costs. Missing within the *GIZ-Perspective* cost are most prominently gross wage costs incurred by the partners employing WAOs and CCs. To these wage costs are added items that have to do with the non-accounting variable costs incurred during the running of the FFBS, that is, the calculatory cost of land, labor, offset by the value of tubers produced on the demo plot. In sum, we then arrive at the *Total Development and Implementation* cost of the FFBS season-by-season, which is compounded up to present time value and aggregated, to yield a cost at present time value of $470,558,952$ KES ($3,913,140$ EUR), or $26,249$ KES (218 EUR) per beneficiary.

	GIZ-Perspective Costs				Wage Costs		Non-Accounting Costs		**Total Dev. & Impl. Costs**	
Year of Training	*No. of Staff-Led FFBS*	*No. of Farmer-Led FFBS*	Total	**Total at Present Value**	Total	Total at Present Value	Total	Total at Present Value	Total	**Total at Present Value**
2016 B	*0*	*0*	3,147,435	**7,338,681**	200,291	467,006	0	0	3,347,725	**7,805,687**
2017 A	*19*	*0*	9,066,304	**19,573,471**	2,144,291	4,629,363	1,093,461	2,360,700	12,304,056	**26,563,534**
2017 B	*18*	*57*	17,320,140	**34,623,041**	3,628,800	7,253,988	3,197,589	6,391,995	24,146,529	**48,269,024**
2018 A	*62*	*54*	27,077,406	**50,118,388**	6,656,727	12,321,138	5,616,042	10,394,901	39,350,175	**72,834,427**
2018 B	*18*	*132*	25,030,710	**42,898,238**	5,019,055	8,601,777	6,041,903	10,354,760	36,091,667	**61,854,775**
2019 A	*55*	*62*	26,512,995	**42,072,791**	6,837,382	10,850,066	5,516,581	8,754,121	38,866,958	**61,676,978**
2019 B	*44*	*116*	27,700,559	**40,701,210**	6,872,727	10,098,291	6,931,431	10,184,546	41,504,717	**60,984,047**
2020 A	*0*	*0*	0	**0**	0	0	0	0	0	**0**
2020 B	*76*	*0*	22,121,754	**27,867,038**	6,939,491	8,741,760	4,373,844	5,509,784	33,435,088	**42,118,582**
2021 A	*0*	*91*	13,782,052	**16,075,386**	3,271,418	3,815,782	3,451,101	4,025,364	20,504,571	**23,916,532**
2021 B	*51*	*0*	13,653,985	**14,746,304**	4,323,927	4,669,841	2,935,079	3,169,886	20,912,992	**22,586,031**
2022 A	*65*	*71*	28,395,784	**28,395,784**	7,120,145	7,120,145	6,433,405	6,433,405	41,949,334	**41,949,334**
				324,410,332						**470,558,952**

Table 12: Aggregating GIZ-Perspective Costs and Total Costs (in KES, except where *italic*) across Seasons

7 The Net Economic Impact of the FFBS

7.1 The Net Social Surplus and Benefit-Cost Ratio

Now that three relevant notions of cost have been produced, we can calculate the net economic impact Π as the difference between gross economic impact and cost, and the benefit-cost ratio π as the ratio of the two, in line with Equations 2 and 3 respectively. Table 13 shows the results. Depending on the cost perspective taken, the value created by the FFBS for the benefit of its beneficiaries net of costs lies between 1.4 and 1.65 billion KES, or 11.8 and 13.6 million EUR. Alternatively, every KES or EUR invested generates between 4 and 7.4 KES or EUR in value.

		KES	EUR
Benefits at Present Value	Gross Economic Impact	1,891,074,413	15,726,060
Costs at Present Value	Partner Perspective	254,788,348	2,118,804
	GIZ Perspective	324,410,332	2,697,777
	Total	470,558,952	3,913,140
Net Economic Impact Π	Partner Perspective	**1,636,286,065**	**13,607,255**
	GIZ Perspective	**1,566,664,081**	**13,028,283**
	Total	**1,420,515,461**	**11,812,920**
Benefit-Cost Ratio π	Partner Perspective	**7.4**	
	GIZ Perspective	**5.8**	
	Total	**4**	

Table 13: The Net Economic Impact of the FFBS

7.2 The Internal Rate of Return

An additional perspective on the net economic impact of the FFBS is the program's seasonal internal rate of return (IRR). The seasonal IRR is defined as the seasonal rate r at which the present time value of benefits equals the present time value of costs. It represents a way to value the program and compare it with other uses for the costs incurred; if the IRR of the FFBS exceeds the IRR of the other uses, then, going by purely financial criteria, the investment in the FFBS is to be preferred over potential other investments.

As shown in Figure 10, the seasonal IRR for the FFBS amounts to 53% if the *Total Implementation and Development* cost is taken to be the relevant cost perspective; it amounts to 70% and 85% for the *GIZ-* and *Partner-Perspective* respectively. For any of the three cost perspectives, the corresponding annual IRR would exceed 100%, which indicates tremendous social return on investment.

A further advantage of this perspective is that it allows the reader to gauge the robustness of the net economic impact figure and the benefit-cost ratio against a range of choices for the seasonal rate r. This rate was calibrated at 8%; However, this calibration is not derived from data or theory beyond a cursory glance at lending rates on uncollateralized credit to smallhold farmers, and therefore merits further scrutiny: Do our results change if a different r is chosen? To our relief, Figure 10 shows that for a range of sensible parameters r (say, for example, a seasonal r between 5% and 20%), the present time values of benefits and costs move nearly in parallel, indicating that the net economic impact Π would be nearly unaffected by a different choice for r. Additionally, since the curves are relatively flat, the benefit-cost ratio π would be only slightly affected. In any case, the FFBS would remain socially profitable up to the intersection of benefit and cost present time values; that is, exactly at the IRRs.

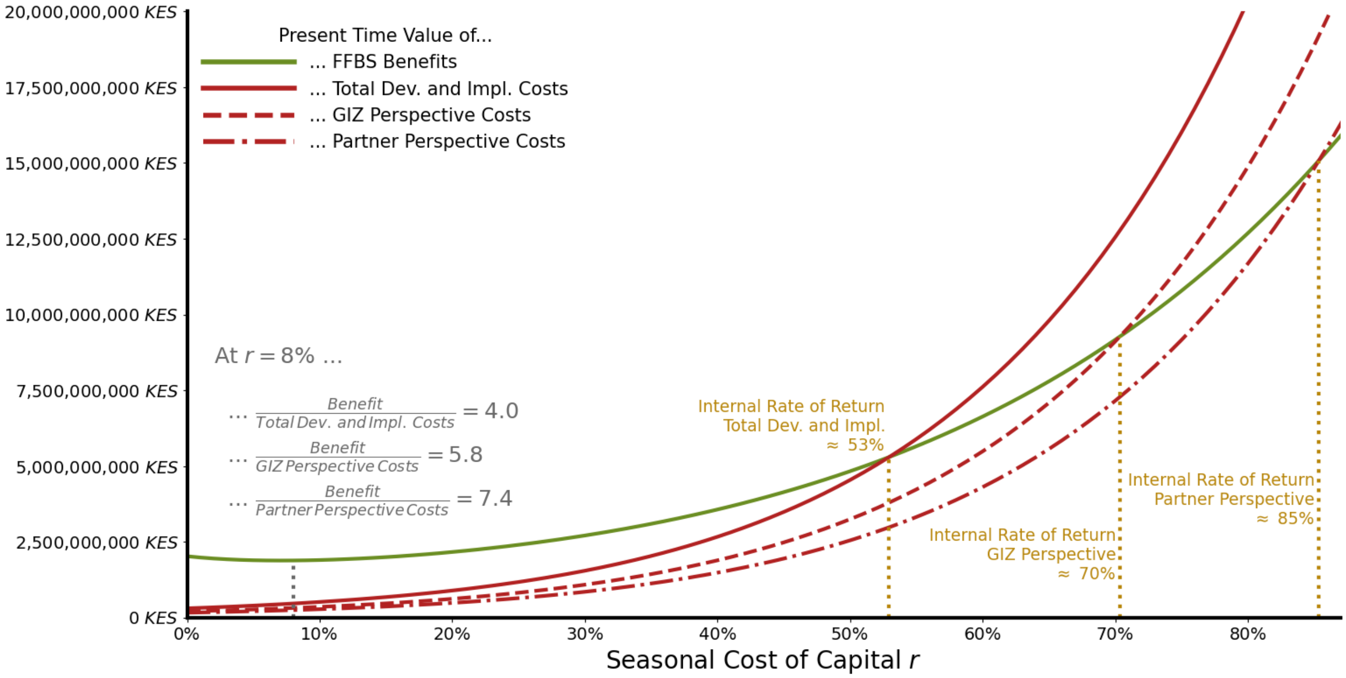

Figure 10: The Internal Rate of Return of the FFBS

8 Conclusion

The goal of the study at hand was to derive the real monetary value created by the FFBS program for the benefit of its participating farmers, and compare it with the costs incurred in developing and implementing the FFBS. To this end, we have built a theoretical framework to derive the economic benefits of the FFBS program using econometric methods. In view of the data required to implement this framework, we have designed a comprehensive questionnaire on tuber yields and prices, agricultural practices, labor costs, demographic information, attitudes and behaviours. We have carefully stratified the sample of respondents according to the population of FFBS-trained farmers, and optimised sampling between the *Treatment* group of trained farmers and the *Control* group of non-FFBS trained farmers in order to achieve an adequate level of statistical power at the lowest possible cost. After successful administration of the questionnaire, we have established a matching scheme to allow a comparison between trained and untrained farmers less encumbered by selection bias, which we further mitigate by focusing on the direct effects of the FFBS. We have built a multivariate model of agronomic value creation and cost, to primarily be able to attribute to the FFBS a gross economic impact, but also allow inference about optimal seed and fertilizer quantities, the perfect planting depth, or the best height at which to hill up. In deriving the cost of the program, we have taken different perspectives on cost that reflect the relevance to different stakeholders, and aggregated all accounted expenditures and non-accounting items to a cost figure commensurable with the gross economic impact, allowing us to find the net economic impact of the FFBS program.

At every step of the way, the aim of the researcher was to derive the net economic impact *conservatively*, that is, to include in the final figure only those drivers of impact which are attributable to the FFBS beyond reasonable statistical doubt; to derive it *empirically*, that is, only on the basis of collected data; to derive it *rigorously*, that is, by employing adequate statistical methods correctly and transparently; and to derive it *comprehensively*, that is, to attempt to capture all drivers of impact via which the FFBS can potentially affect a beneficiary's livelihood, and to include all costs, be they accounted for or not.

Even under such conservativeness and rigor, this study has shown that the FFBS program is highly socially profitable. Over the course of its implementation between 2016 and 2022, it has generated real, economic value of around 1.9 billion KES (15.7 million EUR) for the benefit of 17,927 farmers trained, at a cost between 250 and 500 million KES (2 to 4 million EUR), depending on the perspective taken. Every unit of currency invested generates between 4 and 7.4 units of currency in tangible economic returns.

The social surplus created by the FFBS is mainly driven by knowledge imparted in the domains of soil fertility management and harvest & post-harvest practices, and, to a lesser extent, seed selection and pest & disease management. Additionally, group cohesion is a major driver of social surplus: The staggering persistence of FFBS groups even after training activities have ended is shown in this study to prevent the loss of GAP-related knowledge acquired during the FFBS, thereby contributing more than 200 million KES (1.75 million EUR) in economic impact. Despite some interesting partial findings, the economic impact of the nutrition components of the integrated FFBS could not be conclusively demonstrated, though this is mainly due to the focus of this study on economic impact, quantifiable in terms of currency. Nevertheless, we believe to have shown beyond reasonable doubt that, even when focusing only on the economically quantifiable benefits that are demonstrable under conservative assumptions, the FFBS has a substantial impact on the lives of its beneficiaries that far exceeds the costs of its implementation and development.

References

Arkes, H. R., & Blumer, C. (1985, February). The psychology of sunk cost. *Organizational Behavior and Human Decision Processes*, *35*(1), 124–140. Retrieved from https://doi.org/10.1016/0749-5978(85)90049-4 doi: 10.1016/0749-5978(85)90049-4

Austin, P. C. (2016, August). Variance estimation when using inverse probability of treatment weighting (IPTW) with survival analysis. *Statistics in Medicine*, *35*(30), 5642–5655. Retrieved from https://doi.org/10.1002/sim.7084 doi: 10.1002/sim.7084

Cohen, J., & Dupas, P. (2010). Free distribution or cost-sharing? evidence from a randomized malaria prevention experiment. *The Quarterly Journal of Economics*, 1–45.

Kennedy, G., Ballard, T., & Dop, M. (2013). Guidelines for measuring household and individual dietary diversity. *FAO Reports*.

Riley, J. G. (2001, June). Silver signals: Twenty-five years of screening and signaling. *Journal of Economic Literature*, *39*(2), 432–478. Retrieved from https://doi.org/10.1257/jel.39.2.432 doi: 10.1257/jel.39.2.432

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41–55.

Vagliano, G. (2019). The economic impact of the FFBS program in Bungoma and Nyandarua, Kenya. *GIZ Internal*.

Appendices

A Sampling

Any assessment of the impact of the FFBS will be founded on the comparison between FFBS-trained and non-FFBS-trained farmers. If the questionnaire was to be randomly administered to potato farmers in the counties where the FFBS is active, the resulting dataset might well contain a sample of FFBS-trained farmers, and non-FFBS-trained farmers. However, by controlling the sizes and composition of these samples, termed *Treatment* and *Control* respectively, the researcher gains external validity and statistical power.

A.1 External Validity

The estimation of the net economic impact of the FFBS is based on data collected by sampling the entire population of farmers. Naturally, this sample will not include every single member of the population. Nevertheless, the researcher would like to calculate an impact on the basis of this sample that is, on average, equal to the impact of the FFBS on the entire population of beneficiaries. This requirement is called *external validity*: It is concerned with the generalizability of sample results to a population of relevant subjects. At the stage of sampling, to foster external validity is to build a representative sample: that is, to match the composition of the sample to the composition of the population along relevant dimensions. In our case, we chose to stratify our sample across counties, and across whether a beneficiary was a member of the pure FFBS or a member of the nutrition-integrated FFBS. As can be seen in Table A1, the proportion of *Treatment* group members ought to match the size of the stratified sub-populations. Note that the stratified sub-populations only include farmers trained up to and including Season 2021A, of which there are 14,028 in total. By sampling only from these farmers, we ensure that farmers had time to implement the GAP for at least two seasons, thus avoiding high adoption driven only by recency of training. Note that the FFBS program began in Season 2017A in Nyandarua and Bungoma, and in 2019A was extended to Elgeyo Marawket and Trans Nzoia, thus explaining the lower proportion of trained farmers in the latter two counties.

		FFBS	int. FFBS
Trained Farmers	Nyandarua	6,936	2,841
	Bungoma	1,857	441
	Trans Nzoia	84	1,330
	Elgeyo Marakwet	277	262
Representative Sampling Proportions	Nyandarua	49%	20%
	Bungoma	13%	3%
	Trans Nzoia	1%	9%
	Elgeyo Marakwet	2%	2%

Table A1: Planned Stratified Sampling for the Treated

A.2 Statistical Power

At this point, the composition of the *Treatment* sample is decided; however, the relative size of the *Control* group, and the overall size of the sample, are yet unclear. These quantities are determined by the optimization of statistical power under a budget constraint. Statistical power is the probability that, given that the *Treatment* and *Control* group do systematically differ in some variable, the researcher will detect this difference at some pre-determined level of significance. Naturally, the statistical power depends on choices and parameters: What is the variable in question, how is it distributed, and what is the level of significance at which the researcher wants to detect a systematic difference?

The variable for which we optimize statistical power is ${}^{Yield}/_{ha}$. Thus, in simpler terms, the statistical power answers the following question: If trained farmers truly have higher ${}^{Yields}/_{ha}$

than untrained farmers, how likely are we to detect it?[22] To answer this question, the researcher needs to know the means and deviations of ${}^{Yields}/_{ha}$ within *Treatment* and *Control*. Luckily, the researcher has available agronomic data collected for a previous impact assessment (Vagliano, 2019) to inform these choices, with the required moments reported in Table A2.

	Treatment	Control
Average	12 ${}^{t}/_{HA}$	9.7 ${}^{t}/_{HA}$
Standard Deviation	5.6 ${}^{t}/_{HA}$	8.5 ${}^{t}/_{HA}$

Table A2: The Sample Moments of ${}^{Yields}/_{ha}$ as calculated in 2019

Now, given a size of *Control* and of *Treatment*, a power can be calculated, and visualised as in Figure 11. Green indicates higher power, and red indicates lower power. The graph on the left assumes that the true size of the treatment effect (TE) corresponds to the effect of $12 - 9.7 = 2.3{}^{tons}/_{ha}$ measured in Vagliano (2019), whereas the graph to the right corresponds to a more pessimistic assumption, in which the true treatment effect is assumed to be one standard deviation below the effect measured in Vagliano (2019), at only $0.6{}^{tons}/_{ha}$. Notably, the power depends very sensitively on changes in this true effect size, termed θ. To visualise this dependence further, we plot the threshold at which a power of 80% is reached for different assumptions with respect to θ on the left panel of Figure 12.

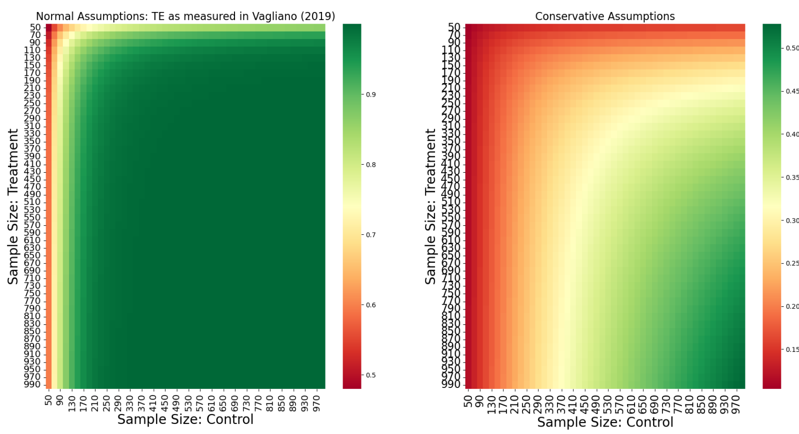

Figure 11: Statistical power of a test for difference in ${}^{Yields}/_{ha}$ at different assumptions

At this point, the keen reader notes that any increase in sample size of *Treatment* or *Control*, the level of power increases, or equivalently that increases in sample size allow the researcher to guarantee a power of 80% at more conservative assumed levels of θ. Since more is always better, there is no scope for optimization yet. However, we now factor in that surveying respondents is costly, and we have limited funds. The optimization problem under a budget constraint can now be posed: it asks us to find the optimal size of *Treatment* and *Control* given a budget allocated to conducting the survey. The budget constraint is expressed as follows:

$$\textit{Man-Days} = \textit{Enumerators} \times \textit{Days} = \frac{N_{Treatment}}{q_{Treatment}} + \frac{N_{Control}}{q_{Control}} \tag{7}$$

where N denotes the sample size and q denotes the number of surveys per enumerator per day that can be carried out for either *Treatment* or *Control* farmers. Assuming $q_{Treatment} = 4$ and $q_{Control} = 2$, which reflects the fact that non-FFBS-trained farmers are much harder to reach than FFBS-trained farmers, for whom contact information is available, we visualise the set of maximally surveyable farmers given different budgets on the right panel of Figure 12.

[22] At a significance level of $\alpha = 5\%$.

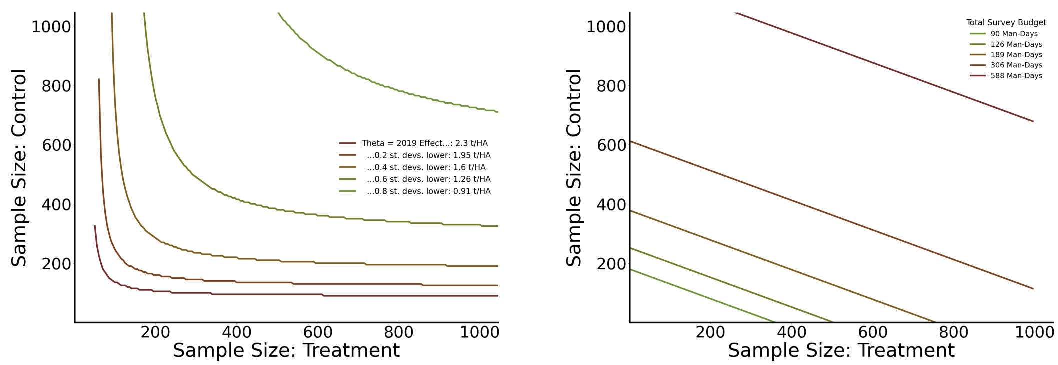

Figure 12: Left: 80% power thresholds at different θ; Right: Budget boundaries

It can be shown that given a certain budget, the greatest level of power achievable corresponds to the power level of a threshold curve tangential to the budget constraint; equivalently, the cheapest way to achieve a given level of power is to select a budget boundary that is tangential to the power curve. In any case, the optimal solution is unique and corresponds exactly to the point at which budget boundary and power threshold touch. The tangentiality condition is visualised in Figure 13, on the left panel. It can also be shown that all optimal points will lie on a line radiating out from the origin; and that every point on the line will constitute a solution for the optimisation problem given some budget constraint. This "optimality line", pictured in grey on the right panel of Figure 13, essentially encodes an optimal ratio of *Control*-to-*Treatment* group members. In our case, this ratio is approximately 1.1, that is, we want 11 *Control* group members for 10 *Treatment* group members. Any sampling that conforms to this ratio is optimal.[23] It remains to choose an absolute level of respondents. This is done either by setting a fixed budget, or by setting a desired θ at which we wish to retain a power level of 80%. In this study, we aim for a power of at least 80% if the true effect of the FFBS is not more than half a standard deviation lower than measured in 2019.[24] This is optimally achieved by a *Treatment* of roughly 320, and *Control* of roughly 350 members, for a total of 670 respondents, resulting in roughly 255 man-days of survey work. However, the sampling outcome is subject to sizeable operational uncertainty, so that any outcome within the "adequacy zone" marked orange-green on the right panel of Figure 13 would be deemed acceptable. To additionally ensure a sampling outcome within this zone, we add a 10% contingency to the labor input, resulting in 280 man-days, and a total survey cost of $1,434,040$ KES, or $1,907$ KES per respondent.

B The Matching Model

In comparing the sample of FFBS-trained (*Treatment*) and non-FFBS-trained (*Control*) farmers directly, the researcher invites selection bias, which can confound inferential results whenever selection into either sample group is non-random. In our case, beneficiaries self-select into *Treatment* (participation in the FFBS) and *Control* (non-participation). Whenever a factor that influences self-selection also influences the level of our variable of interest, bias can arise. Let for example our variable of interest be the propensity to adopt a certain GAP. A possible driver of selection bias might here be well-connectedness of a farmer to her peers: A farmer who is well connected is more likely to join the FFBS (e.g. due to word-of-mouth), but is also more likely to have heard of GAP from her fellow farmers, and therefore already have a higher

[23] The reader observes that even though it is twice as costly to interview a *Control* group member than a *Treatment* group member, we desire a larger *Control*. This is because the variance of $^{Yields}/_{ha}$ among non-FFBS-trained farmers is larger than among FFBS-trained farmers (see Table A2). This implies that the information gained from surveying an additional *Control* group member is more valuable than that of a *Treatment* group member. The optimal solution essentially balances the greater information gain with the higher relative cost.

[24] In this case, the true effect of the FFBS would be not lower than $1.42^{tons}/_{ha}$ in increased yields, instead of the $2.3^{tons}/_{ha}$ measured in 2019. Given the comparatively unfavourable climatic conditions in the 2022A season, for which harvest data was collected, this seems reasonably careful.

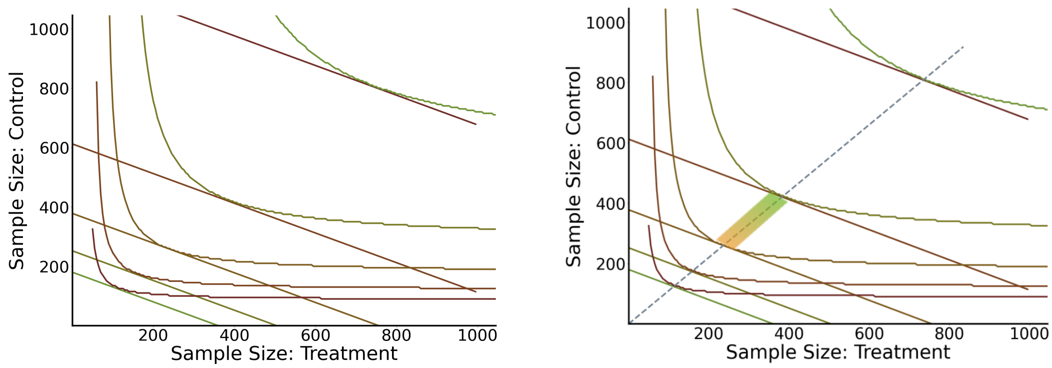

Figure 13: Tangential solutions, the "optimality line" and the "adequacy zone"

degree of GAP adoption. Similarly, a more educated farmer likely has more appetite for further education, and therefore is likelier to select into the FFBS, but will also more likely already have heard of GAP, and be applying them.[25]

To tackle selection bias, we institute a matching scheme. The idea behind matching observations between *Treatment* and *Control* is to approximate a best possible counterfactual dataset, that is, the dataset hidden within the observational data that can be handled as-if it was created by random assignment. Having this subset, the effect of selection into *Control* and *Treatment* samples is said to be *ignorable*. To achieve this, the researcher collects all factors that might influence both self-selection and GAP adoption, and only compares individuals with similar levels of these factors.

B.1 Model Specification

A common candidate for a model that matches observations in *Treatment* and *Control* across multiple dimensions is propensity score matching. In propensity score matching, all characteristics one wants to match across are aggregated into a single score, called the propensity score π. It expresses the probability of selecting into the *Treatment* given a farmer's characteristics. Let $z_i \in \{Control, Treatment\}$ denote a binary variable indicating farmer i's membership in either sample group, and $\boldsymbol{X}_i$ the set of all his or her characteristics. The propensity score can then be expressed as:

$$\pi(\boldsymbol{x}) = P(z_i = Treatment | x = \boldsymbol{X}_i) \tag{8}$$

Naturally, the true probability of selection into *Treatment* is unknown; however, we can estimate it using maximum-likelihood methods. The specification that we choose for our estimator $\widehat{\pi}(\boldsymbol{x})$ is a logistic regression:

$$\widehat{\pi}(\boldsymbol{x}) = \frac{1}{1 + e^{\beta_0 + \beta_1 x_1 + \cdots + \beta_N x_N}} \tag{9}$$

Such a specification has some desirable properties; it is parsimonious, numerically cheap to estimate, and bounded between zero and one, so that any set of farmer characteristics is aggregated into a bona fide probability. Estimation via maximum-likelihood methods yields the unknown β parameters, which, given the correctness of our model specification,[26] are optimal in the sense that they would be likeliest to create the observed split into *Treatment* and *Control*.

As we will see in section B.2 below, propensity score matching is only as good as the estimate for $\pi(\boldsymbol{x})$, which, in turn, is only as good as the choice of its constituent variables $\boldsymbol{x}$. Variables

[25] On the other hand, a farmer with more agronomic education might also have less incentive to join the FFBS due to already knowing the material. In any case, both of these hypotheses can be tested, and should be controlled for when assessing the impact of FFBS on GAP adoption. This can be done by differentiating between the current level of education and the appetite for further education in the questionnaire, and letting these variables enter the estimation of the propensity score, for example.

[26] See footnote 32 above.

ought to be included if they induce selection bias. The extent of the selection bias caused by the omission of a variable is a function of *imbalance* and *importance*. To re-hash an earlier example: well-connected farmers are more likely to be part of FFBS (due to hearing about it via word-of mouth). This is the component of *imbalance*: our variable "well-connectedness to peers" causes *imbalanced* selection into *Treatment* and *Control*. However, *imbalance*, on its own, is not enough to cause selection bias issues. Additionally, a variable has to be *important*, in the sense that it also causes heterogeneous GAP adoption. In the case of "well-connectedness to peers", it does stand to reason that well connected farmers are more likely to have learnt of GAP from their peers, and therefore be stronger adopters. It is the task of the researcher to identify as many of these variables as possible, and include them in $\widehat{\pi}(\boldsymbol{x})$. The final choice of variables entering the estimation can be gleaned from Table A8, along with the estimated parameters.

B.2 The Treatment Effect

Once $\widehat{\pi}$ is estimated, we can calculate an average treatment effect. For any given member i of the population, and for any variable of interest y, the treatment effect is defined as the difference in the level of y between the (potentially hypothetical) situation in which i is assigned to *Treatment*, and the (potentially hypothetical) situation in which i is assigned to *Control*. In our case, the individual treatment effect denotes the change in the level of a variable causally attributable to the FFBS. Sadly for us, this individual treatment effect is unknown: The farmer either selected into *Control* or *Treatment*, so that y is only known for the former or the latter case. However, the average treatment effect across an entire population can be estimated by comparing average values of y in valid *Control* and *Treatment* groups. In a setting where assignment to either sample group is entirely random, simply subtracting the sample average in *Treatment* from the sample average in *Control* yields an unbiased estimator of the average treatment effect. In a setting with self-selection, comparing a propensity-score-matched *Control* with a *Treatment* group approximates random sample assignment, so that a treatment effect across the sample can again be calculated. Before proceeding to show how, two salient points ought to be made.

First, it should be noted that there are two principal ways to aggregate individual treatment effects. On the one hand, the average treatment effect (ATE) is simply the average of all these differences across the entire population. On the other hand, the average treatment effect among the treated (ATT) is the average of all these differences across the population of the treated. In a setting with random sampling, both are, in expectation, the same; in a setting with self-selection, both typically differ. The effect that interests us is the ATT: After all, the economic impact of the FFBS arises in the population of FFBS-trained farmers. To calculate the impact within this population, we are interested in the causal impact of the FFBS on a certain variable y for a typical trained farmer, not for any farmer.

Second, it should be noted that in Section 4.2, in equation 5, we expressed how to calculate the individual partial impacts γ^e, that is, the surplus generated by the FFBS for the benefit of its beneficiaries, mediated via effect variable e. For any effect variable e, this equation relates the economic surplus created by the implementation of e with the increase in the adoption of e directly attributable to the FFBS. We now see that this latter increase, termed Δ^e, corresponds definitionally to the ATT for e.

Hence, an estimator for the ATT yields the final missing piece in estimating gross individual impacts, and ultimately the net economic impact causally attributable to the FFBS. We can then proceed with the calculation of this estimator for the ATT, termed $\widehat{\Delta}^e$, which is constructed in propensity score matching as follows. Instead of simply subtracting sample means of *Control* and *Treatment* from each other, we subtract a weighted mean of the *Control* sample from the simple mean of the *Treatment* sample.

$$\widehat{\Delta}^e = \mu_T^e - \mu_C^e \tag{10}$$

where, letting Z_i^e denote that value associated with effect variable e and farmer i:[27]

$$\mu_T^e = \frac{\sum_{i \in T} Z_i^e}{|T|} \tag{11}$$

[27]that is, if e is top dressing, then, if farmer i top dresses, $Z_i^e = 1$, else $Z_i^e = 0$.

and

$$\mu_C^e = \frac{\sum_{i \in C} w_i Z_i^e}{\sum_{i \in C} w_i} \tag{12}$$

where the weight of each observation reflects the predicted odds of self-selecting into *Treatment*.

$$w_i = \frac{\widehat{\pi}(\boldsymbol{X}_i)}{1 - \widehat{\pi}(\boldsymbol{X}_i)} \tag{13}$$

It can be shown that an estimate for an average treatment effect using these weights is unbiased in the presence of selection bias in sample assignment (Austin, 2016). Intuitively, a member of *Control* who is judged by the propensity model to have been very likely to select into *Treatment* is a more valuable point of comparison for members of the *Treatment* group, and therefore receives a higher weight, than a member of *Control* who is judged to be unlikely to have selected into *Treatment*. If one thinks of the propensity model output, the probability, as a measure of "similarity with *Treatment*", then propensity score matching can be thought of as giving a higher weight to members of *Control* who are very similar to members of *Treatment*. The weighted *Control* outcomes can then be thought of as coming from a *Matched Control* group, which is as similar as possible to the *Treatment* group, but did not select into the FFBS. Hence, comparison of *Treatment* with *Matched Control* allows the researcher to approximate the effect in variable e directly attributable to the FFBS, without the influence of group composition effects that result from self-selection into our sample groups.

C The Revenue Model

The revenue model R_i^e is likely the most crucial component of individual partial impacts γ_i^e. Recall that R_i^e denotes the additional revenue gained by implementing effect variable e. In the binary case, we compare the predicted revenue given e is implemented, and subtract from it the predicted revenue given that e is not implemented, all else being equal.[28] These revenues are calculated, as one would expect, by multiplying quantity and unit value. Usually, unit value corresponds exactly to the price a kilogram of tubers fetches on the market. However, not all potatoes sold fetch the same price: Seed potatoes are likely to be more expensive than ware potatoes. Also note that some tubers, such as potatoes kept as seed or consumed, create value without having an explicit price. However, we still want to include their value in the revenue calculation, even though the revenue here is only indirectly realized. After all, tubers kept as seed do not have to be bought from certified seed producers or other farmers. Hence, they have economic value in that they allow the farmer to save on seed-related expenses. Each of the following classes $\boldsymbol{K}$ will have a different unit value calculated for it:

1. Ware (Sold)
2. Seed (Sold)
3. Seed (Kept)
4. Derived Product (sold)
5. Consumed
6. Given Away
7. Rotten/Fed to livestock/Other

[28] Note that we use a model prediction for both possible values of the effect variable, even though we will actually observe one of these revenues in the data. However, for the sake of consistency, we will use predicted values for both. In the aggregate, this will likely make very little difference since our revenue model is an estimator for expected revenue conditional on regressor values. Hence, on aggregate, an average across these expected conditional revenues should lie close to average observed conditional revenue.

For every class $k \in \boldsymbol{K}$, we compute a revenue by multiplying the class quantity function $f^k \times Y$ with a class unit value function V^k. Note that the class quantity function is in turn calculated by multiplying f^k, the function giving the proportional size of a class in %, with a total yield function Y. The product evaluated at $e = 1$ is subtracted from the product evaluated at $e = 0$ to yield R_i^e. All other function inputs are taken to correspond to the characteristics X of farmer i sourced from the data collected.

$$R_i^e = \sum_{k \in \boldsymbol{K}} \Big[f^k(\boldsymbol{x} = X_i | e = 1) \times Y(\boldsymbol{x} = X_i | e = 1) \times V^k(\boldsymbol{x} = X_i | e = 1) \\ - f^k(\boldsymbol{x} = X_i | e = 0) \times Y(\boldsymbol{x} = X_i | e = 0) \times V^k(\boldsymbol{x} = X_i | e = 0) \Big] \tag{14}$$

It remains to understand where functions f^k, Y, and V^k come from. A function is a transformation rule that tells us how a collection of inputs is transformed into a single output. To clarify this by way of example: Y is a rule that tells us how much a farmer harvests per hectare, given all relevant factors (farmer experience, application of GAP, labour inputs, weather, soil quality, etc.). These factors are the input, yield per hectare is the output of the function. Sadly for us, the general rule by which yield per hectare is determined is unknown; even if it were known, the rule would likely be much too complex to express in mathematical form. Hence, the researcher has to approximate the rule; to this end, he specifies and estimates a model.

C.1 Model Specification

The specification outlines the structure of the model; it determines how the input factors combine to give a total yield per hectare. In our case, the model specification looks like this:

$$\frac{Yield}{Ha} \equiv Y = \beta_0 \times \underbrace{A_{\mathrm{GAP}_1} \times \cdots \times A_{\mathrm{GAP}_N} \times A_{\mathrm{IDDS}}}_{\textit{Effects of GAP and Food Diversity}} \times \underbrace{C_1 \times \cdots \times C_M}_{\textit{Controls}} \times \left(\frac{Labor}{Ha}\right)^{\beta_1} \tag{15}$$

Depending on the definitions of regressors of interest A and control regressors C, such a specification corresponds to a Cobb-Douglas production function. Such a production function has useful characteristics for our application. For example, factor inputs are complementary: As the amount of land increases, labour has to increase also, to sustain a given yield per hectare. At no labour input, yield is zero. Every additional unit of labour adds a diminishing amount of yield per hectare (if $0 < \beta_1 < 1$): This is the property of diminishing returns. The impact of each GAP is neatly captured in the expressions A_{GAP}, which can take different forms. For a binary effect variable such as Top Dressing, we have:

$$A_{TopDressing} = \begin{cases} 1 & \text{if top dressing is not applied,} \\ 1 + \beta_{TopDressing} & \text{if top dressing is applied.} \end{cases} \tag{16}$$

which implies that implementing top dressing will increase $\frac{Yield}{Ha}$ by approximately $\beta_{TopDressing} \times 100\%$.[29] For non-binary effect variables such as seed density, the A expression can look a bit more intricate

$$A_{Seed} = \left(\frac{\text{Seed}}{\text{Ha}}\right)^{\beta_{Seed}} \tag{17}$$

but still has an elegant interpretation[30] and the same desirable properties that the Cobb-Douglas production function affords: Complementarity of inputs, no yield if the seed rate is zero, and diminishing returns.

Similarly, for the unit value of sold ware and seed potatoes, we specify a price model:

$$V^k = P^k = \gamma_0 \times \underbrace{B_{\mathrm{GAP}_1} \times \cdots \times B_{\mathrm{GAP}_N}}_{\textit{Effects of GAP on Price}} \times \underbrace{C_1 \times \cdots \times C_M}_{\textit{Controls}} \times \underbrace{\left(1 + \gamma_{Seed} \mathbb{1}_{k=Seed\ (Sold)}\right)}_{\textit{Seed Potato Mark-Up}} \tag{18}$$

[29] See Footnote 14.

[30] increasing seed rate by 1% increases yields by $\beta_{Seed}\%$

where the multiplicative Cobb-Douglas specification allows us to easily include a mark-up factor on seed potatoes relative to ware potatoes, in addition to factors relating to the GAP, or control variables. In this way, we are able to show the revenue impact of GAP such as Sorting & Grading, which influence only the price of sold potatoes.

For the other, more marginal classes in $\boldsymbol{K}$, specifying a Cobb-Douglas function is not necessary. Here, we can choose simpler approximations for the price (or, more generally, the economic value of the items of class k). In the case of potato-derived products, their unit value is held constant at the average price observed in the data collected:

$$V^{Derived} = \overline{P_i^{Derived}} \tag{19}$$

The value derived from consumed potatoes is calculated by modelling the diminishing marginal utility of potatoes consumed using a sigmoid function

$$u(q) = \frac{1 + e^{-ks}}{1 + e^{k(y-s)}}$$

which expresses the fraction of utility that the q^{th} kilogram of potatoes consumed carries relative to the first kilogram of potato consumed.[31] $u(q)$ can then be integrated between zero and the total quantity of potato consumed Q to find the cumulative utility (relative to the first kilogram consumed) afforded by the total level of potato consumption per household per season:

$$U(Q) = \int_0^Q u(q)dq \tag{20}$$

Letting the utility of the first kilogram consumed be equal to the average price of potatoes sold as ware, we can calculate the absolute utility afforded by the consumption of harvested potatoes, and dividing by the total number of potatos consumed, we find the average value of a harvested tuber consumed by the household.

$$\overline{U}(Q) = \frac{U(Q)}{Q} \times \overline{P_i^{Ware\ (Sold)}} \tag{21}$$

Evaluating this function at the predicted level of consumed ware potatos generates a value function $V^{Consumed}$ that depends on characteristics x:

$$V^{Consumed}(x) = \overline{U}((f^{Consumed} \times Y)(x)) \tag{22}$$

Potatoes which were given away, rotten, fed to livestock, or other, are assigned a value of zero.

$$V^{Given\ Away}(x) = V^{Rotten/Fed\ to\ livestock}(x) = 0 \tag{23}$$

The proportion according to which harvested potatoes fall into one of the classes k is specified in function f^k. Since f^k designates a proportion, the sum over all classes has to yield 1; in other words, the total harvest should be completely accounted for when adding up all its possible uses. A popular specification for such a function expressing a proportion is the softmax, according to which:

$$f^k = \frac{e^{\delta_0^k + \delta_1^k x_1 + \cdots + \delta_N^k x_N}}{\sum_{k \in \boldsymbol{K}} e^{\delta_0^k + \delta_1^k x_1 + \cdots + \delta_N^k x_N}} \tag{24}$$

Here, the normalization factor in the denominator, which is simply the sum of the numerators across all categories, ensures that the proportions sum up to 1. The numerator determines the

[31] Here, k (set to 0.03) is a parameter expressing the smoothness of the sigmoid, whereas s encodes the saturation point, that is, the quantity of potatoes at which the marginal utility (that is, the utility of the next kilogram consumed) is exactly half of the utility of the first kilogram consumed. If the daily saturation point per person is assumed to be two medium sized potatos, the seasonal saturation point per household comes out to

$$s = 2 \times {}^{Days}/_{Season} \times {}^{People}/_{Household} \times {}^{Weight}/_{Tuber}$$

for which estimated values can be plugged in:

$$s = 2 \times 180 \times 3.8 \times 0.2 \text{ kg} = 273,6 \text{ kg}$$

share of yield going towards k. It consists of an exponentiated linear function $e^{\delta_0^k+\delta_1^k x_1+\cdots+\delta_N^k x_N}$ in which different inputs x_1 to x_N are allowed to influence the share of yields going towards k, weighted by parameters δ_1^k to δ_N^k respectively. The inputs x are sourced from the data collected from farmers. They include, for example, the perceived access to market that a farmer has. If a farmer perceives to be struggling to sell her potatoes, then her share of ware sold is likely to be lower and her share of wastage is likely to be higher than a farmer who perceives to be able to easily find buyers for her tubers. Hence, perceived market access is a good input x for the modeller to choose, among others.

C.2 Model Estimation

So far, we have outlined how we intend to *specify* the functions Y, V^k, and f^k that constitute our revenue model, that is, we have outlined the general shape that we impose on them. However, all these functions still have unknown components. These are the parameters β, γ and δ, which occur in Y, V^k, and f^k respectively. Not knowing these components prevents us from making actual predictions about yield, price and revenue. Thankfully, we can estimate these parameters using the survey data that has been collected. For such an estimation to be possible, both the function inputs and the output must be known; that is, we need to know the total yield harvested, and all function inputs that we wish to include in the specification for total yield Y; we need to know the prices of goods sold, and all input factors that we wish to include in the specification for prices P_k; we need to know the shares of goods going towards class k, and all factors that we wish to include in the specification for proportions f_k. Thankfully, the survey has been designed with these needs in mind, so that we can estimate parameters β, γ and δ for all input variables that we deem relevant.

To estimate these parameters, we employ Maximum-Likelihood methods. Intuitively, these methods yield those parameters that have the greatest likelihood of generating the data observed, given a model specification. In other words, since we have specified general transformation rules Y, V^k, and f^k already, and we have collected data corresponding to the input and the output variables of these models, we can use Maximum-Likelihood methods to find those parameters which best represent the data. They are those parameters that, if the general transformation rules were true,[32] and if we were given all values of the input variables, would be likeliest to create the observed values of yield, price and proportion.

For the yield and price models, estimation proceeds as follows. First, the multiplicative specification is log-linearized by taking the natural logarithm ln on both sides of the equation. This takes us from here...

$$\frac{Yield}{Ha} = \beta_0 \times A_{\mathrm{GAP}_1} \times \cdots \times C_M \times \left(\frac{Labor}{Ha}\right)^{\beta_1} \tag{25}$$

... to here:

$$ln(Yield) - ln(Ha) = ln(\beta_0) + ln(A_{\mathrm{GAP}_1}) + \cdots + ln(C_M) + \beta_1 \times ln(Labor) \tag{26}$$

This equation can now be estimated using so called Ordinary Least Squares (OLS). OLS finds the maximum-likelihood parameter values by minimizing the sum of squared residuals, i.e. the squared differences between predicted model values and the observed output values. For the yield proportion models, we employ multinomial logistic regression to find the maximum-likelihood parameter values δ^k.

The interpretation of these parameters depends strongly on the model specification. Consider for example our yield model Y, which is multiplicative in its inputs. In such a specification, the estimated parameters usually indicate how a unit change in input affects the output in percentage terms. For example, our $\beta_{TopDressing}$, the parameter associated with top dressing in the yield model, is equal to 0.1865. This means that the model believes that, at all other values held constant, a farmer who was previously not doing top dressing and then decides to implement it can expect an increase in yields of around 18.7%, approximately.[33] In the yield

[32] In the sense that they do correctly combine all knowable inputs to a prediction of output, and any deviations of observed output to predicted output are due to truly random error.

[33] See Footnote 14.

proportion functions f_k, interpretation is trickier. Here, a unit change in variable x_n implies a δ_n^k change in the log-odds of yield class k relative to a baseline class.[34] Intuitively, one can say that if a parameter δ_n^k in equation k is comparatively large, an increase in the value of variable n associated can be said to increase the proportion of tubers going towards class k.

D From Seasonal Impact to Lifetime Impact

Before proceeding to calculate the gross economic impact, the reader notes that the individual impact γ calculated as per Equation 1 is a seasonal impact, that is, it represents the FFBS-attributable value per hectare that a single farmer reaped over the single season between April and August 2022, termed 2022A. However, note that this FFBS-attributable value recurs over many seasons, starting from the season after which the farmer was trained.[35] Hence, the seasonal impact has to be aggregated up to a lifetime impact, which is incurred over all seasons within a farmer's lifetime after training.

When performing such an aggregation, one commonly calculates a present time value, that is, the value of past and future returns, expressed in today's units of currency. Returns incurred at different time points can not simply be added up: The researcher must account for the fact that one KES earned in a season from now has slightly lower value than one KES earned today.[36] The rate at which this nominally identical amount of currency differs in real terms is called the discount rate, which we will derive in Appendix D.1. Armed with this discount rate, we then calculate a multiplier that transforms past and future seasonal impacts into one lifetime impact in Appendix D.2.

D.1 Discount and Compound Rates

The discount rate broadly consists of two elements: p and r. On the one hand, p is the probability that the advantage conferred by better agricultural know-how garnered during FFBS training is lost from one season to the next. The potential reasons for this loss of advantage are many-fold. For example, if knowledge about a GAP is forgotten, or a specific GAP becomes obsolete, then from that point on, future partial returns deriving from this GAP do not occur. Similarly, if a GAP that is learned about in the FFBS would have been learned about anyway a few seasons later (due to other sources of agronomic learning), then the knowledge of the GAP is said to be supplanted, and the impact attributed to the FFBS from this specific GAP also has to be assumed to fall to zero. This is not because the GAP is not being applied anymore, but because the GAP is being applied regardless of FFBS participation from that point on. In other words, the knowledge imparted by the FFBS is only relevant for those first few seasons where knowledge about the GAP would not have been acquired anywhere else. Alternatively, a farmer might quit farming altogether; if this happens, the agricultural know-how may persist, but does not confer any economic benefits. The joint probability of all these factors (forgetting, obsoletion, or supplantment of GAP, quitting farming, etc.) is termed p. This factor p serves to discount future returns in expectation, since they only occur with probability $1-p$ next season. In turn, because past returns have already been realized, p is not applicable to past seasons.[37]

[34] In our case, the class of tubers sold as ware.

[35] To err on the side of conservativeness, that is, to under- rather than overestimate, benefits are assumed not to be incurred in the season the farmer was trained. One could make the case that this season should be included, since the FFBS training occurs in real-time, so to speak: Agricultural practices are demonstrated on the demo plot at the same time they ought to be applied on a farmer's own plot, as per the GAP. Hence, some benefits of GAP application are conferred in the season of training itself. However, the farmer receives the demo plot's harvest as seed, which can only be planted in the season after training. For the GAP concerned with seed quality, only the season after training can be assumed to have adoption figures representative of a trained farmer. Therefore, to be safe, economic value created by the farmer is only attributable to the FFBS beginning from the season after training, for *all* GAP and other effect variables.

[36] Within the scope of our study, "today", or "present time", denotes the end of Season 2022A, that is, the end of August 2022. Usually, the calculation of a present value involves only discounting, and the calculation of a future value only compounding. In our case, where all monetary quantities are expressed in currency as per the end of August 2022, we use the term *Present Time Value* to denote the value of a stream of benefits or expenditures as of that moment in time.

[37] One could argue that at least some elements of p could serve to compound past returns. For example, since GAP adoption rates are measured in the survey at least two seasons after training has been completed, one could

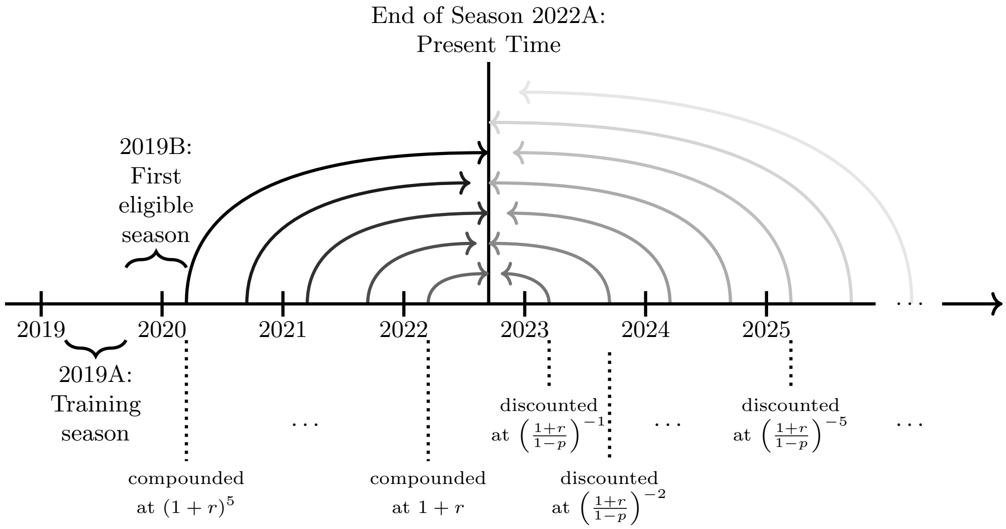

Figure 14: Calculating the present value of impact for a farmer trained in Season 2019A

On the other hand, r quantifies by how much one KES earned today is worth more than a KES earned next season. This difference in value arises because a KES earned today can be put into a savings account, or, if the farmer is debt financed, does not have to be borrowed, so that the farmer either earns a savings rate or saves on a lending rate. As yet another alternative, the farmer could have invested the KES earned today in productive assets (such as equipment, machinery or land), where it would have accrued a return during the next season that would not have accrued had the investment been made a season later. Between these three, the typical scenario, from anecdotal evidence, is that a smallhold farmer is debt financed, and so a lending rate for private persons without collateral should be taken as the benchmark r. A KES earned today is worth $1+r$ KES tomorrow; inversely, the value of one KES earned in a season from now, measured in today's terms, is $\frac{1}{1+r}$. As opposed to p, r is applied to future and past returns equally. Combining both p and r, future returns are discounted at the rate $\frac{1+r}{1-p}$, whereas past returns are compounded at $1+r$. Figure 14 outlines the discounting/compounding schedule visually.

D.2 The Season-to-Lifetime Multiplier

We can now calculate a multiplier M that converts seasonal to lifetime impact as the sum of the present time value of past returns and future returns, which, since discount and compound rates differ, are calculated separately.

$$M = M_{past} + M_{future} \tag{27}$$

where

$$M_{past} = 1 + (1+r) + (1+r)^2 + \cdots + (1+r)^T = \frac{(1+r)^T - 1}{r} \tag{28}$$

with T denoting the number of seasons passed since the end of training, and

$$M_{future} = \frac{1-p}{1+r} + \left(\frac{1-p}{1+r}\right)^2 + \left(\frac{1-p}{1+r}\right)^3 + \left(\frac{1-p}{1+r}\right)^4 + \ldots = \frac{1-p}{r+p} \tag{29}$$

argue that, in the past seasons, GAP adoption rates were even higher than measured here, and were subsequently partially forgotten, fell into disuse otherwise, were supplanted, and so on. These higher adoption rates would then have caused a higher seasonal impact in past seasons. However, for the sake of conservativeness, we will not appreciate past seasons by $1-p$ per season, instead cautiously assuming that the current level of adoption represents the level of adoption at the end of training. See more on the extent and impact of knowledge decay in Appendix E.

Age Bracket	Proportion	Probability of Retirement
15-19	1%	0%
20-29	6%	0%
30-39	18%	0%
40-49	27%	1%
50-59	29%	10%
60-69	15%	20%
70-79	4%	25%
≥ 80	0%	30%

Table A3: Age brackets, their representation, and their assumed probability of retirement

exploiting the convergence of geometric series.[38]

It now remains to assign values to p and r. An attempt is made in Appendix E to empirically quantify the degree to which GAP adoption decays as seasons pass, either due to forgetting of practices or willful non-application. Additionally, a toy model allows a rough estimate of the proportion of farmers retiring each year, which likely is a major driver of FFBS benefits lost due to farmer exit. This toy model assigns each surveyed farmer an annual probability to retire. Assuming regular retirement at 60, members in the age bracket 50-59 are assigned a 10% chance to retire; older cohorts are assigned a higher chance to retire, younger cohorts a lower chance. The exact probabilities are reported in Table A3. Using the proportional representation of age brackets and the probability of retirement, we can calculate that the mean annual chance to retire stands at 7.5%. Assuming that about half of all retirees pass their farms to relatives who were taught agricultural practices by the retiring farmer (thereby passing on the knowledge gained by the FFBS), the real loss of GAP-related knowledge, and therefore the loss of the FFBS benefits, amounts to just under 4% annually, or 2% seasonally. Naturally, retirement is not the only factor driving farm exit; the researcher ought to account for job changes (within or without the agricultural sector, but away from potato farming), injury or death. Especially the rate of turnover might not be negligible, however, no data could be found on this. Therefore, the researcher assumes a 3% seasonal probability of exit due to all factors except retirement, yielding (just under) a 5% seasonal probability of the loss of FFBS benefits due to farm exit.

The probability that knowledge becomes obsolete is not estimable from our data alone; it depends on the future development of agricultural research and innovation. The probability that knowledge acquired in the FFBS becomes supplanted by knowledge that would have been gained anyway had the farmer not participated in the FFBS is similarly inscrutable: an estimate would depend on the structure of agricultural knowledge dispersion. Hence, the researcher has to make an educated guess at these probabilities. For the probability of obsoletion of GAP-related knowledge, the researcher assumes 5% seasonally. This would imply that, after 10 years, only around 35% of the FFBS curriculum is still up-to-date, representing a quite pessimistic estimate. Hence, the estimate is reasonably conservative - note that a higher rate of obsoletion serves to devalue future returns in present currency terms, hence lowering the gross economic impact of the FFBS program. For the probability of supplantment and the probability of a farmer dropping out of potato farming between seasons, the researcher assumes 5% each. Given that, in Appendix E, we show that there is no indication of adoption decay due to forgetting of practices or willful non-application, we have all the components needed to calculate the joint probability p. It comes out to $1 - 0.95^3 = 14.2\%$, which we round up to an even 15%.

Orienting one on the typical uncollateralized lending rate offered by Kenyan banks and Saccos to private persons of around 1.25% p.m, the resulting seasonal lending rate, which serves as a good proxy for r, amounts to around 8%. Plugging in p and r into the formula for M_{future}, we obtain 3.7, whereas the average multiplier M_{past} is 6.[39] Adding both components up, we obtain a total multiplier of $M = 9.7$.

[38] Note that we assume an infinite remaining farmer lifespan for ease of computation. Since, at sensible p of around 15%, and a sensible interest rate of around 5%, knowledge depreciates very quickly - later seasons carry rapidly diminishing value - this assumption is very innocent.

[39] The average farmer was trained 4.75 seasons ago. However, note that by Jensen's inequality (see Footnote 17), plugging in T = 4.75 into Equation 28 does not yield the correct, average multiplier.

E Adoption Decay and Group Cohesion

In our gross economic impact calculation, parameter p represents the seasonal probability that agricultural knowledge gained in the FFBS becomes forgotten, supplanted or obsolete, or is otherwise not applied anymore. Jointly with the rate of capital depreciation r, parameter p determines M_{past} and M_{future} as in Equations 28 and 29. These two components make up M, which translates the seasonal gross economic impact into a lifetime impact. As such, the result of this assessment is highly dependent on the values chosen for p and r. While a qualified estimate for parameter r can be made using interest rate data, parameter p is more elusive. In this section, we will attempt to estimate a sub-component of this parameter, termed p', and at the same time assess the value of group cohesion, that is, the value of remaining in the FFBS group after training has ended.

E.1 Estimating p'

As stated earlier, p represents the probability that, between seasons, the knowledge on the GAP gained from prior FFBS training is forgotten, is willfully not being applied, is supplanted by knowledge acquired regardless of training, becomes fully obsolete, or is redundant due to farmer exit caused by for example retirement, change of jobs, or other factors. Though some of these probabilities are unquantifiable by nature (obsoletion, supplantment) or due to lack of data (farmer turnover). However, using the data at hand we can give an estimate for the rate of agricultural knowledge being forgotten, or otherwise not being applied anymore. This rate of adoption decay, which we will term p', will represent a lower bound for parameter p, since it does not account for supplantment, obsoletion, or farmer turnover.

To estimate p', we attempt to link an individual FFBS-trained farmer's rate of GAP adoption to the time elapsed since her training. The modelling framework is as such: Let t_i denote the number of seasons elapsed since training of farmer i, and GAP_{i,t_i} denote the average level of GAP adoption of i at that time. At $t_i = 0$ specifically, $GAP_{i,0}$ denotes the average level of GAP adoption during the season of FFBS training. Note that FFBS training occurs in real-time, so to speak: Agricultural practices are demonstrated on the demo plot at the same time they ought to be applied on a farmer's own plot, as per the GAP. Hence, during the season of training, no adoption decay can have occurred. Every season after completion of training, the level of adoption decays toward the level of an untrained farmer, $\overline{GAP}$, at rate p'. In our case, this level of adoption is taken to be the average rate of adoption of a farmer in the *Matched Control* group, which stands at 53%. This decay dynamic can be written concisely as in Equation 30, where the current level of adoption by farmer i is represented as a weighted mixture between the initial level of adoption at the time of training and the *Matched Control* level of adoption, with the weights determined by the rate of decay p' and the time elapsed since training t_i.

$$GAP_{i,t_i} = GAP_{i,0}\left(1-p'\right)^{t_i} + \overline{GAP}\left(1-\left(1-p'\right)^{t_i}\right) \tag{30}$$

Factoring and log-linearizing, we arrive at

$$ln\left(GAP_{i,t_i} - \overline{GAP}\right) = \underbrace{ln\left(GAP_{i,0} - \overline{GAP}\right)}_{\beta_0} + t_i \underbrace{ln\left(1-p'\right)}_{\beta_1} \tag{31}$$

Note that the left-hand-side of this expression is known: it represents the difference between the rate of GAP adoption of a trained farmer and the *Matched Control* level of adoption.[40] We also know, for each farmer, the season of her training, so that t_i can be derived. Hence, the linear Equation 31 can be estimated empirically, by ordinary least squares, resulting in coefficients β_0 and β_1.

Running this regression on the *Treatment* sample data results in a positive parameter $\beta_1 = 0.011$. Note that, as per the curly bracket in Equation 31, a positive parameter β_1 actually

[40] Note that the decay model as stated in Equation 30 only makes sense if $GAP_{i,t_i} > \overline{GAP}$. Hence, for the sake of the estimation at hand, *Treatment* adoption rates are bounded below by $\overline{GAP}$, plus a small ϵ. Empirically, this bounding is only moderately relevant; out of 365 *Treatment* farmers, only 28 have an adoption figure lower than $\overline{GAP}$, the mean of *Matched Control*.

implies a negative p'. In this case, GAP adoption does not decay, it actually increases as time goes by. However, since our parameter is not different from zero with any statistical significance, all that can be stated is that the underlying data does not show any signs of GAP adoption decay over time. Visually, this is illustrated in Figure 15, where the solid grey line represents the best OLS fit. Indeed, it slopes very slightly upwards, implying a non-significantly negative p'.

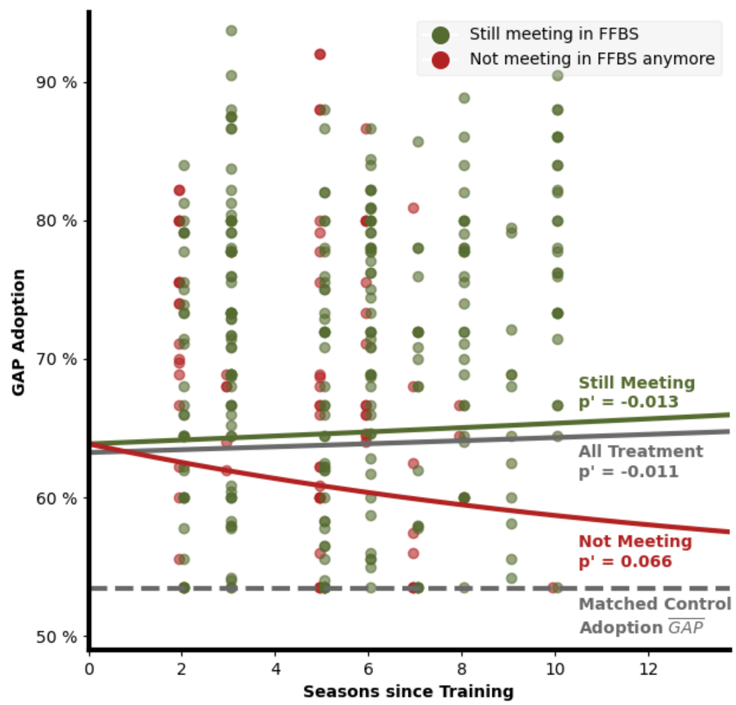

Figure 15: The Decay of GAP Adoption for Subgroups of the *Treatment* sample.

E.2 The Value of Group Cohesion

The simplicity of the model setup above allows us to test whether continued group membership can counteract adoption decay. To this end, we can extend the linearized model outlined in Equation 31 to predict a separate β, and thereby a separate p', for those farmers still meeting with their FFBS group, and those not meeting anymore. These p' are also visualised in Figure 15. Pictured in green, those still meeting with their FFBS group show a rate of adoption decay not significantly different from zero. On the other hand, those not meeting with their group anymore, pictured in red, show a significantly positive rate of decay. For them, p' is estimated to be around 0.066, indicating that, every season, 6.6% of remaining knowledge gained in the FFBS on GAP adoption is either forgotten or is not applied for other reasons.

To be able to quantify the economic impact of remaining active within one's FFBS group, we can now use the calculated decay parameter $p' = 6.6\%$ for a trained farmer who is not active within her group to calculate a present value multiplier. This is done as in Equation 29, except that an additional 6.6% are tacked on to $p = 15\%$. This can be compared to the present value multiplier of a trained farmer who still meets in her group, for whom p' is assumed to be zero. Under these assumptions, the multiplier for a non-meeting farmer stands at 8.6, whereas it stands at 9.7 for a farmer that still meets her FFBS group. In other words, a farmer still meeting her group enjoys the benefits conferred by the FFBS for just over a full season more, measured at present value. Stated in economic terms, around 10% of the overall gross economic impact of calculated in Section 5.2 is attributable to remaining an active member within one's FFBS group.

F Nutrition

A potential vector of economic impact is the nutritional training that members of the nutrition-integrated FFBS participate in. In addition to the GAP, the nutrition-integrated FFBS training goes through a curriculum concerned with dietary diversity, hygiene, food preparation, preservation and production, and child nutrition. Along these areas of focus, multiple potential pathways are imaginable for nutritional knowledge gained to have a quantifiable economic impact. The study at hand focuses on three possible pathways: The impact of dietary diversity on productivity, the connection between dietary diversity and the propensity to rotate crops, and the effect of nutritional training on the extent of value-addition measures taken.

F.1 Dietary Diversity and Productivity

The most immediate way in which nutritional training could potentially have an impact on farmer surpluses is by impacting productivity. An increase in productivity attributable to the training would show up in the agronomic models of yield or price as a regressor. As with the impact of the agricultural and farm-business-related FFBS training, however, the impact of the nutrition component needs to be mediated via an effect variable. Our study focuses on the possible mediation via dietary diversity. In effect, we ask whether the nutritional training, by increasing participating farmer's dietary diversity (and thereby influencing beneficiary health, labor days available and productive capacity) also improves the yields harvested. Our measure of dietary diversity, which will be included in the regression model as an effect variable, is the Individual Dietary Diversity Score (IDDS).

The IDDS is a survey-based measure of dietary diversity developed by the FAO (Kennedy, Ballard, & Dop, 2013). The respondent is asked to recall all types of food and drink consumed in the last 24 hours. For each of a total of 16 categories, the respondent scores a point if he or she ate at least one food belonging to this category. Some of these categories are aggregated, and others are omitted, to yield a total of nine categories,[41] each scored either zero or one. The IDDS relates the sum of these binary scores, and therefore ranges between zero and nine.

As seen in Table A4, the IDDS achieved by *Control* stands at almost exactly 5. For the sample weighted to match the *Treatment*, the IDDS is moderately higher, at 5.2. Compared to this figure, the IDDS increases by .3 in the pure FFBS sample, and .4 in the nutrition-integrated FFBS sample. This effect is driven by results in Nyandarua, where the IDDS for *Treatment* is markedly higher than in *Control* and *Matched Control*. By contrast, in the western counties, the IDDS of trained farmers is lower than that of non-FFBS-trained farmers, but insignificantly so.

	Control	Matched Control	FFBS	Δ	int. FFBS	Δ
Nyandarua	4.98	5.11	5.75	+.63 ***	5.94	+.83 ***
West	5.18	5.45	4.92	-.53	5.05	-.40
All	**5.03**	**5.22**	**5.53**	**+.31** **	**5.65**	**+.43** **

Difference between *Treatment* and *Matched Control* highlighted.
$*: p < 0.05$, $**: p < 0.01$, $***: p < 0.001$

Table A4: The Individual Dietary Diversity Score (0-9) across the Sample

In summary, there is statistically significant evidence that trained farmers, and nutrition-integrated farmers especially, score higher in the IDDS against their non-FFBS-trained counterparts. However, to establish the economic value of this increased level of dietary diversity, dietary diversity has to be shown to influence farmer yields. Unfortunately, including the IDDS in a multitude of specifications for the yield model never results in a coefficient remotely close to statistical significance. Hence, a direct economic impact of the nutrition integrated FFBS via dietary diversity can not be established.

[41] to wit: Starchy Staples; Fruit and Vegetables Rich in Vitamin A; Other Fruit and Vegetables; Dark Green Leafy Vegetables; Meat and Fish; Organ Meat; Eggs; Legumes, Nuts and Seeds; Milk and Milk Products

F.2 Dietary Diversity and Crop Rotation

A more indirect way by which the nutrition-integrated FFBS may cause increases in farmer surplus might be established via the link between dietary diversity and crop rotation. One might reason, for example, that sensitization to the importance of dietary diversity might encourage crop rotation. After all, potato is a starchy staple that is widely consumed across the sample: out of the nine condensed IDDS classes, the propensity to consume starchy staples is highest, standing at 98%. Hence, a preference for a more varied diet, as induced by the nutrition-integrated FFBS, might encourage rotating crops away from potato after having planted potato, thus promoting dietary diversity via the adoption of a good agricultural practice, crop rotation.

There is some evidence that crop rotation does indeed increase dietary diversity. Table A5 shows results for a variety of estimated models attempting to find the drivers of food diversity. The rightmost column shows the impact of these drivers on the IDDS directly. As we can see, if potato has been planted twice consecutively on the main plot, the IDDS decreases by slightly more than 0.4 points, controlling for preferences in food diversity, participation in the pure or nutrition-integrated FFBS, and the size of the main plot relative to all arable land. This latter regressor is interesting in and of itself: it stands to reason that, if the main plot takes up a large proportion of all land, food diversity is reduced, since little else can be grown except the one crop planted on the main plot. Indeed, the associated coefficient is negative, albeit insignificantly so.

While the direct effect of crop rotation on the IDDS is interesting, it does not tell a very nuanced story. The underlying mechanism hypothesized by the researcher argues that repeatedly planting potatoes on the main plot increases the propensity to consume potatoes, but decreases the propensity to consume other, less prevalent food classes. Since starchy staples are widely consumed anyway, the increase in potato consumption does little to increase the IDDS, whereas the reduction in consumption of other food classes decreases the IDDS.

This mechanism can indeed be tested: We can build models predicting the probability to consume a food class given the same regressors as above. The model estimates are reported in the first ten columns of Table A5. Note that we have split up the category of starchy staples into the two constituent food groups: cereals and white root tubers. The latter contains potatoes, so we will take a look at it first. As hypothesized, if potato is planted consecutively, the probability of consuming white root tubers does indeed increase, as evidenced by the significantly positive coefficient. On the other hand, most other food groups see a reduction in consumption probability if potato is planted twice consecutively. Most saliently, the propensity to consume fruit and vegetables is reduced drastically. This also is consistent with the hypothesis formulated above.

Despite these hopeful signs, a positive impact of the nutrition-integrated FFBS on dietary diversity via crop rotation can not be established. This is because participation in the nutrition-integrated FFBS does not increase the propensity to rotate crops beyond the effect of the pure FFBS. As Table A6 shows, there is no indication that knowledge about the importance of dietary diversity fosters crop rotation beyond what is taught in the pure agricultural curriculum.

This same reason also explains why the nutrition-integrated FFBS can not be shown to have an impact on yields via crop rotation. Additionally, recall that the estimation of the effect of crop rotation on yields is confounded by innate soil fertility (see section 5.1.9). Hence, it is impossible to establish an economic impact of the nutrition component of the nutrition-integrated FFBS mediated via crop rotation.

F.3 Value-Addition Measures

As part of the nutritional curriculum taught in the nutrition-integrated FFBS, farmers learn about measures that can be taken to add value to harvested potatoes by turning them into crisps or potato cakes, and selling them. In fact, these measures can potentially increase revenues quite substantially: among those that were post-processing potatoes for added value, the average revenue gained amounted to almost $25,000^{KES}/_{ha}$, per season. However, the prevalence of these measures is marginal, in the low single digits. Hence, the average revenue gained per hectare from these measures, even though indeed highest in the nutrition-integrated FFBS, is negligible.

Probability of consuming...	... Cereals	... White Root Tuber	... Fruit and Vegetables Rich in Vitamin A	... Other Fruit and Vegetables	... Dark Green Leafy Vegetables	... Meat and Fish	... Organ Meat	... Eggs	... Legumes, Nuts and Seeds	... Milk and Milk Products	**IDDS**
Constant	3.5458*** (1.0473)	1.5263*** (0.5528)	0.5764 (0.4187)	2.1678*** (0.6440)	2.4205*** (0.7470)	-2.1302*** (0.6236)	-2.0167*** (0.5697)	-0.8182* (0.4352)	1.4526*** (0.5110)	-0.3855 (0.4085)	**5.2075*** (0.2806)**
Potato planted consecutively	-0.2038 (0.5696)	0.5690* (0.3009)	-0.7502*** (0.2406)	-1.2092*** (0.2981)	0.0560 (0.3394)	-0.4092 (0.2939)	0.4638 (0.3830)	-0.2739 (0.2619)	0.4329 (0.3261)	-0.3287 (0.2442)	**-0.4180** (0.1641)**
Size of main plot rel. to all arable land	-1.3593** (0.5554)	0.0021 (0.2841)	-0.5885** (0.2591)	-0.3158 (0.4107)	-0.3488 (0.3640)	0.1553 (0.2781)	0.7839* (0.4142)	0.0385 (0.2622)	-0.7134** (0.2988)	0.0540 (0.2516)	**-0.2652 (0.1714)**
Participated in FFBS	0.1882 (0.4331)	0.3606* (0.1921)	0.3996** (0.1810)	1.3218*** (0.3974)	0.4390* (0.2654)	0.3859** (0.1913)	0.0499 (0.3165)	-0.0037 (0.1825)	0.4697** (0.2217)	0.2965* (0.1741)	**0.4529*** (0.1183)**
Participated in nutr. integr. FFBS	0.6750 (0.6417)	0.7083*** (0.2570)	0.7586*** (0.2388)	0.4210 (0.3723)	1.0644** (0.4186)	0.1888 (0.2440)	0.1980 (0.3771)	0.2280 (0.2228)	0.6192** (0.2934)	0.4340** (0.2175)	**0.6213*** (0.1477)**
I need a little variety in my diet once in a while	0.5746 (1.1478)	-0.4236 (0.5838)	0.3792 (0.4480)	-0.2337 (0.6715)	-1.0262 (0.7690)	1.0762* (0.6432)	-1.1109* (0.6423)	0.0176 (0.4621)	0.1621 (0.5480)	0.2823 (0.4328)	**0.0887 (0.2979)**
I need some variety in my diet regularly	0.0539 (1.0608)	-0.7388 (0.5603)	0.1872 (0.4264)	0.2485 (0.6609)	-0.3994 (0.7594)	0.9762 (0.6271)	-0.6922 (0.5793)	0.1487 (0.4408)	-0.0167 (0.5198)	0.0471 (0.4141)	**0.1103 (0.2848)**
I need a lot of variety in my diet every day	-0.1152 (1.1114)	-1.5979*** (0.5749)	-0.6163 (0.4470)	-0.0930 (0.6991)	-0.5733 (0.7877)	0.5594 (0.6515)	-1.2848* (0.6753)	-0.2778 (0.4691)	-0.0212 (0.5490)	-0.8651* (0.4459)	**-0.5137* (0.3002)**

Table A5: Modelling the Probability of Consuming a Food belonging to an IDDS Class, and the IDDS itself

			Control	Matched Control	FFBS	int. FFBS
21	Non-Consecutive Potato	Potato was not planted in the same field in which potatoes were planted in the current season.	82%	87%	95%	94%
22	Recommended Rotation	The crops planted before, during and after the season in question all come from a different family of plants.	44%	51%	70%	61%

Table A6: The Adoption of GAPs related to Crop Rotation, by FFBS type

	Control	Matched Control	FFBS	int. FFBS
Propensity to perform value addition	2%	1%	0%	3%
Revenue/ha from value addition	536 KES	210 KES	0 KES	945 KES

Table A7: The Prevalence and Impact of Value Addition Measures

G Model Regressions

Explanatory Variables $\boldsymbol{X}$	Probability to select into *Treatment* π
Constant	-2.6935*** (0.6350)
Age	0.0195** (0.0076)
Experience in Potato Production	0.0253** (0.0110)
Sex: Male	-0.6499*** (0.1721)
Level of Education (0: No Schooling, 6: Post-Graduate)	-0.0897 (0.1133)
Interest in Learning About Agronomics (0: Not interested at all, 3: Very interested)	0.3755*** (0.1430)
Total Land in Acres	-0.0335 (0.0242)
Preference for food diversity (0: I am content eating the same foods every day, 3: I need a lot of variety in my diet every day)	0.5860*** (0.1162)
Perceived Demand for Ware Tubers (0: I struggle to sell what I harvest, 3: I could sell much more than I harvest)	0.3249*** (0.0813)
Source of Agronomic Info: Media	-0.0268 (0.1734)
Source of Agronomic Info: Informally, through peers	-1.4532*** (0.1712)
Source of Agronomic Info: Input suppliers	0.1798 (0.1953)
N	752
Pseudo-R^2	0.1489

Table A8: Regression for the Propensity Score

	Explanatory Variables $\boldsymbol{X}$	$log(^{Yield}/_{Ha})$	$log(^{Price}/_{Kg})$	$log\left(\frac{f(Consumed)}{f(Sold\ as\ Ware)}\right)$	$log\left(\frac{f(Sold\ as\ Seed)}{f(Sold\ as\ Ware)}\right)$	$log\left(\frac{f(Kept\ as\ Seed)}{f(Sold\ as\ Ware)}\right)$	$log\left(\frac{f(Given\ Away)}{f(Sold\ as\ Ware)}\right)$	$log\left(\frac{f(Rotten)}{f(Sold\ as\ Ware)}\right)$
Controls	Constant	-0.9971*	0.6357***					
		(0.5900)	(0.1755)					
	$log(^{Yield}/_{Ha})$			-0.1867***	-0.368***	-0.172***	-0.307***	-0.303***
				(0.0366)	(0.0624)	(0.0382)	(0.0638)	(0.0694)
	$log(^{Yield}/_{Ha})$ in Vicinity	0.5508***						
		(0.0645)						
	$log(^{Price}/_{Kg})$ in Vicinity		0.7981***					
			(0.0503)					
	Land was Fallow	-0.0460						
		(0.0823)						
	Land was Virgin	0.2596*						
		(0.1400)						
	Age of Farmer	-0.0044*						
		(0.0024)						
	Experience of Farmer	0.0053						
		(0.0035)						
	Sex of Farmer	-0.0844						
		(0.0525)						
	Land Owned	0.0076						
		(0.0078)						
	Variety: Not Shangi	0.2015**						
		(0.0895)						
	Variety: Mix or Unknown	0.0907						
		(0.1026)						
	Potato not planted consec.	-0.2686***						
		(0.0866)						
	Disease Infestation Severity	-0.0869**						
		(0.0406)						
	Pest Infestation Severity	-0.1065***						
		(0.0300)						
	Sold at Farm Gate		-0.0332	-0.5943*	-0.0969	-0.3435	-0.6245*	-0.3012
			(0.0352)	(0.2666)	(0.3282)	(0.2611)	(0.3666)	(0.4629)
	Sold to Customers			2.049***	0.7406	0.9359***	1.5214***	1.1767***
				(0.2258)	(0.7024)	(0.2556)	(0.3678)	(0.3738)
	Sold to Middle Men		-0.0914**					
			(0.0385)					
	Sold to Processor		0.0462					
			(0.1747)					
	Sold to Trader		-0.0197					
			(0.0455)					
Seed Selection	Positively Selected	0.1214						
		(0.0857)						
	Clean/Certified	0.1247*						
		(0.0703)						
	$log(^{Seed}/_{Ha})$	0.4843***						
		(0.0401)						
Soil Fertility Management	$log(^{Plant.\ Fertilizer}/_{Ha})$	0.1688***						
		(0.0320)						
	$log(^{Top\ Fertilizer}/_{Ha})$	0.0529**						
		(0.0259)						
	Applied Recs. of Soil Test	0.1865						
		(0.1168)						
	Applied Manure	-0.0931						
		(0.1415)						
Land Preparation	Timing	-0.1263*						
		(0.0735)						
	Steps	0.0072						
		(0.0616)						
	$log(Depth)$	0.2125***						
		(0.0664)						
Planting Practices	$Depth$	0.0237						
		(0.0170)						
	$Depth^2$	-0.0012**						
		(0.0006)						
Weeding, Hilling & Thinning	$Height\ of\ Hills$	0.0061						
		(0.0124)						
	$Height\ of\ Hills^2$	-0.0001						
		(0.0003)						
	Thinning Done	0.0205						
		(0.0853)						
Pest and Disease Management	Rogueing Done	0.0299						
		(0.0852)						
	Disease Chemicals Used	0.1594**						
		(0.0716)						
	Pest Chemicals Used	0.0621						
		(0.0661)						
Harvest and Post-Harvesting	Dehaulming Done	0.1037*						
		(0.0617)						
	Sorting & Grading Done		0.0707	0.3333	0.4972	0.6081*	0.5003	0.0281
			(0.0515)	(0.2577)	(0.3584)	(0.2991)	(0.4991)	(0.434)
Marketing Group	Membership		0.0834**					
			(0.0356)					
Summary Statistics	N	751	644	751	751	751	751	751
	R^2	0.5204	0.3081					
	Adjusted R^2	0.4997	0.3005					
	Pseudo-R^2			0.3023	0.3023	0.3023	0.3023	0.3023

Table A9: Regression Models for Yield, Price and Proportions

As a federally owned enterprise, GIZ supports the German Government in achieving its objectives in the field of international cooperation for sustainable development.

Commissioned by:
Deutsche Gesellschaft für
internationale Zusammenarbeit (GIZ) GmbH

Registered offices
Bonn and Eschborn

Friedrich-Ebert-Allee 36 - 40
53113 Bonn, Germany
T +49 228 44 60 - 0
F +49 228 44 60 - 17 66

Dag-Hammarskjöld-Weg 1-5
65760 Eschborn, Germany
T +49 61 97 79 - 0
F +49 61 96 79 - 11 15

E info@giz.de
I www.giz.de

Contact:
Thomas Miethbauer
Adviser and Head of Programme (AV) at Rural Development & Agriculture
Division Value Creation, Innovation and Employment (G510)
Global Programme Promotion of Nutrition Sensitive Potato Value Chains
in East Africa
E thomas.miethbauer@giz.de

Author:
Gianluca Vagliano
T +49 171 6018559
E gl.vagliano@gmail.com

Editors:
Thomas Miethbauer, Festus Kimatu

Photo credits/sources:
© GIZ Jackson Muchoki, Meshack Ronoh

Frankfurt, April 2023

Deutsche Gesellschaft für
Internationale Zusammenarbeit (GIZ) GmbH

Registered offices
Bonn and Eschborn

Friedrich-Ebert-Allee 32 + 36
53113 Bonn, Germany
T +49 228 44 60-0
F +49 228 44 60-17 66

Dag-Hammarskjöld-Weg 1 - 5
65760 Eschborn, Germany
T +49 61 96 79-0
F +49 61 96 79-11 15

E info@giz.de
I www.giz.de