FROM DRAWING TO REALITY:

THE IMPACT-OF THE PIP-APPROACH ON FARMERS' MOTIVATION, RESILIENCE AND STEWARDSHIP IN RURAL BURUNDI

JULY 2020

OXFAM

Table of contents

Appui au Développement Intégral et à la Solidarité sur les Collines

Nos Principes: Savoir, Pouvoir, Avoir, Vouloir, Agir

Nos Valeurs: Ecoute, Pro-Activité, Respect de la Dignité Humaine, Transparence, Performance, Intégrité, Diversité

Developing Agriculture from the Ground Up

from relief to recovery

ACKNOWLEDGEMENTS

This report is based on the information provided by hundreds of people interviewed across Burundi in December 2019 and January 2020. First and foremost, we want to express our gratitude to all who participated in the evaluation. Their willingness to give their time and discuss their knowledge and opinions with the interviewers made this evaluation possible. Overall, this resulted in valuable information for the programme implementation and will serve as endline data for evaluating the impact of the PIP-approach.

This evaluation was achieved through the extensive contributions and expertise of the entire PAPAB consortium IFDC, Oxfam Novib, Wageningen Environmental Research, ZOA, Adisco, OAP and Réseau Burundi 2000+. In particular we would like to thank Micael Beun and Bizabityo Déogratias at IFDC as well as Belyse Rudaragi at Oxfam in Burundi, for their support in fielding the survey, and their continued assistance during data collection and drafting the report. Special thanks go to Koos Michel, Ruben de Winne, and Nynke Kuperus for their insightful reviews of this report. Finally, Aad Kessler at Wageningen Environmental Research deserves special mention here for his helpful reviews, his constructive and critical support and throughout the whole process of conducting the PIP-impact study.

EXECUTIVE SUMMARY

This report presents an impact assessment of the PIP-approach as implemented within the PAPAB-programme in Burundi (Projet d'Appui a la Productivité Agricole au Burundi) between November 2015 till May 2020.

The PAPAB-programme is funded by the Embassy of the Kingdom of The Netherlands in Burundi and aims to sustainably increase food production in Burundi by promoting market-oriented, climate-resilient, and sustainable agricultural techniques, supported by targeted fertilizer subsidies. The PAPAB-programme brings together a consortium of NGO's. This consortium is led by IFDC and has Oxfam Novib, Wageningen Environmental Research, and ZOA as its lead implementing partners, who further engages with three local organizations: Adisco, OAP and Réseau Burundi 2000+.

The PAPAB-programme employs the PIP-approach as iis core intervention strategy in its engagements with farmers throughout Burundi. The PIP-approach acknowledges the importance of collaboration between farmers, and communities and aims to foster resilience-based stewardship through an interrelated set of activities. In the PIP-approach, households jointly devise a vision and plans for the future of their farm. Training, demonstrations and competitions of various (innovative) agricultural practices and techniques are offered in support of the implementation of these PIP-plans. In most cases, training and demonstrations are delivered by peers (other PIP-farmers). Thereby, the first generation, selects the second generation of farmers through demonstrations and competitions, upon which these existing generations of PIP-farmers select and train subsequent generations, exponentially increasing the number of PIP-farmers in each colline. The PIP-approach aims to increase farmers motivation, their resilience, and their stewardship. The objective is to change the mindset of farmers which are "farmers by default" (without alternatives) to gradually become "farmers by choice" who are good stewards of their land and sustainable entrepreneurs.

Using data collected among 962 farmers spread across 35 collines in 5 provinces in Burundi this impact evaluation employs a quasi-experimental design where each generation of PIP-farmers is matched with a similar group of non-PIP farmers who reside in the same agro-ecological zones in Burundi. The study then assesses the PIP-approach's effectiveness on a multifaceted set of outcomes that are structured within the pillars motivation, resilience, and stewardship. Next to an assessment of effectiveness, this study also further studies the theoretical backbone of the approach. By investigating the relationships between motivation, stewardship, and resilience, this study dives deeper into the mechanism and causal pathways that drive farmers to become more resilient and good stewards of their land.

The results from a wide range of rigorous statistical analyses show that the PIP-approach is effective. PIP-farmers are more motivated, more resilient farmers, and better stewards of the land compared to similar farmers that did not take part in the PIP-approach. Moreover, this study shows that that PIP-farmers are reaping benefits of their participation beyond increases in motivation, resilience, and stewardship. In a set of self-assessments, the bulk of PIP-participants reports that their income from agricultural-sources, the number of different crops that they cultivate, and the number of different crops they sell at markets has increased substantially over the past three years. Analyses further demonstrate that PIP-farmers' food security situation is less volatile throughout the year, and they are more food secure, particularly during the lean season compared to non-PIP farmers.

The first generation of farmers is composed of relatively well-to-do farmers, whereas later generations are lower on the socio-economic ladder, particularly generation 3 and generation 4. Our analyses show that the sequential targeting strategy (where generation 1 selects generation 2, and so forth)

used by the project is inclusive in the sense that it does reach poorer farmers in later generations. This is evidence for the mechanism that earlier generations of PIP-farmers lead by example and that their innovative mindset drives later generations of PIP-farmers to participate in the approach as well. Generally, we find the strongest impacts on motivation, resilience, and stewardship of the project among the first generation of farmers (which is also the generation that has participated the longest and has already implemented the majority of their PIP-plans) followed by subsequent and later trained Generations 2, 3, and 4.

Although overall levels of resilience and stewardship are higher among PIP-farmers compared to non-PIP farmers, we do not find an impact on some behavioural aspects of resilience and stewardship. We find no impacts on farmers crop and livestock diversity, particularly among later generations. Regarding stewardship, we find that PIP-farmers implement some farming practices more often compared to non-PIP farmers (especially crop-rotation) and that PIP-farmers are more knowledgeable about a wider range of farming practices. PIP-farmers however, do not implement a markedly distinct set of farming practices compared to their non-PIP peers. Rather, a larger share of PIP-farmers implements those practices that are relatively common among non-PIP farmers as well.

Higher levels of motivation are associated with more resilience and better stewardship of the land. Moreover, we find support for the notion that involving the whole household in making an integrated farm plan is important. The whole household is the cornerstone upon which the PIP-approach is rolled out. The focus on the whole household that jointly develops a vision for the future acts as a catalyst for resilience and stewardship as well. In our analyses, we find that it is especially those aspects of motivation that capture levels of collaboration and support within the household that are important determinants of farmers resilience and also their stewardship. Our analyses corroborate the mechanisms flowing from the theoretical foundation of the approach.

We find that most gains in motivation and resilience are made early on in the project, that is, in the early stages of the implementation of PIP-plans. Motivation, stewardship, and resilience reach a ceiling and marginal gains in motivation, stewardship, and resilience wear-off in later stages of implementing the PIP-plan.

Given the learning that takes place and the high gains in motivation, stewardship, and resilience early on in the programme we recommend a quicker follow up from one PIP-generation to the next compared to the current situation (after a period of 2- 3 years) wherever possible as most gains are made during these early stages of PIP-implementation.

Our results advise against devising a more narrowly defined PIP-approach in the future, for instance, by cherry-picking or choosing to single out a narrow sub-aspect to work when designing in the PIP-activities. Our analyses demonstrate that the holistic and integrated approach as set out in the theoretical framework is what delivers impact.

We recommend to put more emphasis on livestock and crop diversification of PIP-farmers as early on in the PIP-approach as possible. Provide more tools and skills that can help PIP-farmers to diversify the production of both crops and livelihood products. Analyses show that this is not only important for increasing resilience but having a diversified farm is also an important determinant of stewardship. Because we find that PIP-farmers are more knowledgeable about a variety of farming practices but refrain from implementing these, we recommend to carefully study the barriers PIP-farmers face in the implementation of not-so-common farming practices that foster stewardship of their land.

All in all, this study shows that the PIP-approach is very effective in increasing motivation, resilience, and stewardship, its theoretical backbone and assumptions are well-grounded and hold empirically after rigorous testing. Therefore our final recommendation is to scale up the PIP-approach to other suitable areas in Burundi.

1 INTRODUCTION

1.1 THE PAPAB-PROGRAMME

The PAPAB-programme (Projet d'Appui a la Productivité Agricole au Burundi) aims to sustainably increase food production in Burundi by promoting market-oriented, climate-resilient, and sustainable agricultural techniques, supported by targeted fertilizer subsidies. The project is funded by the Embassy of the Kingdom of the Netherlands. The PAPAB consortium responsible for programme delivery is lead by IFDC, with Oxfam Novib, Wageningen Environmental Research, and ZOA as its lead implementing partners Adisco, OAP and Réseau Burundi 2000+. The PAPAB-programme was rolled out between November 2015 and May 2020. The programme's set of activities focus on (a) increasing soil fertility through seeking reforms to fertilizer subsidy systems in Burundi and (b) increasing agricultural productivity, improving farmers' resilience, and increasing farmer's and producer's organisations access to markets.

In its direct engagement with farmers and the activities aimed to increase sustainable food production, farmer's resilience and market access, the PAPAB-programme implements the PIP-approach (Plan Intégré du Paysan) as its core intervention strategy. This impact study focuses on the PIP-approach used as an intervention strategy in the wider PAPAB-programme.

The PIP-approach acknowledges the importance of collaboration between farmers, and strong communities and aims to foster resilience-based stewardship through an interrelated set of activities based on households jointly devising a vision and plans for the future of their farm. This includes training, demonstrations and competitions of various (innovative) agricultural practices and techniques, which are (in most cases) delivered by peers (other PIP-farmers). This impact study focuses on the PIP-aspects of the PAPAB-programme through an assessment of the impact of the PIP-approach on farmers motivation, resilience, and their stewardship of the land and the interplay between these three main pillars of the PIP-approach.

1.2 THE PIP APPROACH

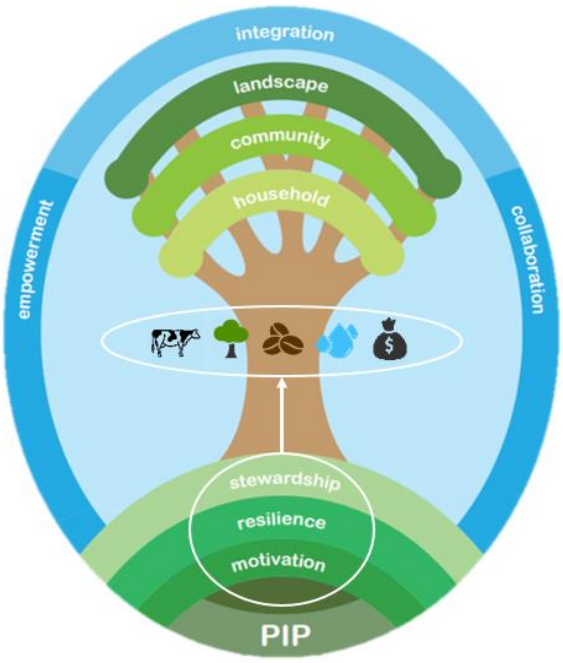


Figure 1 The PIP-approach visualised as a tree: improvements in farmers livelihoods at the household, community, and landscape are rooted in farmers' stewardship, resilience and motivation.

The starting-point of the PIP approach is the notion that resilient farming systems, are grounded in a solid foundation of motivated people and healthy land. The PIP approach builds these foundations. With a dynamic process of vision building, planning, learning and action, the PIP approach generates an ever-increasing number of proud farmers who realize that their land is their main asset, and who feel able and intrinsically motivated to invest in their farms. This is a huge change because once this foundation is laid, collaboration with other stakeholders and integration of new activities drives larger-scale change. This means building the capacity of smallholder farmers and increasing their motivation to experiment with improved practices. Learning from others is then conducive to collective action and the take-up of village-wide resilient farming. The objective is to change the mindset of farmers which are "farmers by default" (without alternatives) to gradually become "farmers by choice" (land stewards, sustainable entrepreneurs).

The PIP approach is illustrated in Figure 1 as a tree that grows in fertile soil. The approach is rooted in the notion of first investing in people and land they manage, before investing in anything else. Motivation, resilience and stewardship of the land are thus seen as a pathway towards sustainable changes in farmers' households, the wider community and ultimately the landscape. If changes are rooted in people's motivation, and these are based on a vision and plan shared by the family, it is more likely that improvements in farmers lives and their communities, changes to the landscape such as reforestation and soil improvements are more successful. This is illustrated by the arrow in the trunk of the tree pointing to activities that aim to improve agronomic conditions, improve productivity and yields and land degradation. Such efforts require genuinely motivated farmers to achieve sustainable outcomes. The scaling-up of the PIP approach and the mobilization of farmers to collaborate is crucial to reverse land degradation. The branches of the tree illustrate this where the process starts at the household level, then spreads to the community level and eventually covers the structural changes in the whole landscape.

The outer circle of Figure 1 presents the three guiding principles of the PIP approach: empowerment, integration and collaboration. These are particularly important during PIP-implementation and show that it is always crucial (especially for project staff) to empower people, to foster the integration of activities, and to enhance collaboration to go to scale faster.

The PIP approach starts with drawing the farm's current situation. Then this drawing is compared to a drawing showing the farm's household ideal vision. Farmers subsequently develop a concrete action plan that describes the (agronomic) practices that need to be implemented to arrive at this ideal envisioned situation (the Integrated Farm Plan or PIP-plan). The discussions surrounding the formulation of the PIP induces enhanced cooperation and social cohesion between family members. Discussions within families shape the plans for improvements in their livelihoods. The PIP-plan transforms the vision into attainable goals. The process of formulating and implementing this plan creates is geared towards more resilient farming, increases awareness about changes in land and soil and deepens understanding about crop and land management practices so PIP-farmers will become good stewards of their land. Figure 2 gives an example of a PIP as drawn by a family, with left the current situation and right the desired future farm in 3-5 years.

Figure 2 An example of an Integrated Farm Plan (current situation on left and desired further on the right-hand side)

1.3 FOUR GENERATIONS OF PIP FARMERS

The PIP approach starts in a village with farmer innovators (PI's, 'Paysans Innovateurs') who spread their knowledge to the next generations of PIP farmers. PIs are trained in various techniques, ranging from crop management to soil erosion practices. The PIP-activities draw specific attention to land management, crop, and livestock practices, however non-farm, household and entrepreneurial activities are included as well. PI's then transfer their (agronomic) knowledge through farmer-to-farmer training and through competitions. The competitions and trainings centre around the creation and implementation of the PIP-plan, and are expected to raise awareness and the farmers' intrinsic motivation throughout the village and beyond, laying the basis for more trust, collaboration and sustainable development within the villages.

The PI's are the first generation of PIP-farmers. The PI's are farmers are jointly selected by other community members (it is made clear that eventually, any farmer in a colline will be able to join in later generations). The PI's are trained in various agronomic techniques by the PAPAB consortium partner organisations across specifically selected collines across Burundi. They reside in so-called 'original collines' (as this is where the roll-out of the PIP-approach started). After a first growing season and having implemented practices planned in their PIP on the farm, the first generation then organizes a competition with demonstrations to train the 2nd generation of farmers. The 1st generation selects the participants for this training (the 2nd generation) predominantly within their own social circle within their own colline. The 2nd generation of farmers is then trained by the 1st generation of farmers.

Subsequently, after the next growing season, the next generation of farmers in the original collines are invited to participate in the PIP-approach, through demonstrations and competitions. These constitute the 3rd generation, and participating farmers are then trained by farmers from both the 1st and 2nd generation during a next PIP competition. During this same 2nd PIP competition, some

farmers from adjacent villages (so-called extension collines) also participate on their own initiative, especially the most interested ones, who then act as pioneers in their own (extension) colline.

Another PIP competition is then organised in the extension collines, and PIP farmers who have already created their PIP in these collines (hence, 3rd generation PIP farmers) are now PIP trainers themselves: the farmers that they train in these extension/adjacent collines during the competition become the 4th generation PIP farmers. This roll-out exponentially increases the number of PIP-farmers in each generation.

2 RESEARCH QUESTIONS

2.1 RESEARCH QUESTIONS

The key questions for this study can be broken down in two strands of research questions. We distinguish between ***assessing effectiveness*** on the relevant outcomes (RQ 1) and unpacking the ***causal chain*** behind the PIP-approach by understanding the relationships between these outcomes (RQ 2).

2.1.1 ASSESSING EFFECTIVENESS OF THE PIP APPROACH

RQ 1: To what extent can differences in motivation, resilience, and stewardship be attributed to the implementation of the PIP-approach among smallholder farmers in Burundi?
When assessing effectiveness (RQ 1), we focus on motivation, stewardship and resilience as our main outcomes. We assess the impact of the approach across the four generations of farmers that participated in the PIP-approach. The main assessment of the project impact lies in comparing whether those farmers that participated in the PIP-approach display higher motivation, stewardship, and resilience, compared to the group of farmers that did not participate in the project (a so-called counterfactual) who are similar to the participants.

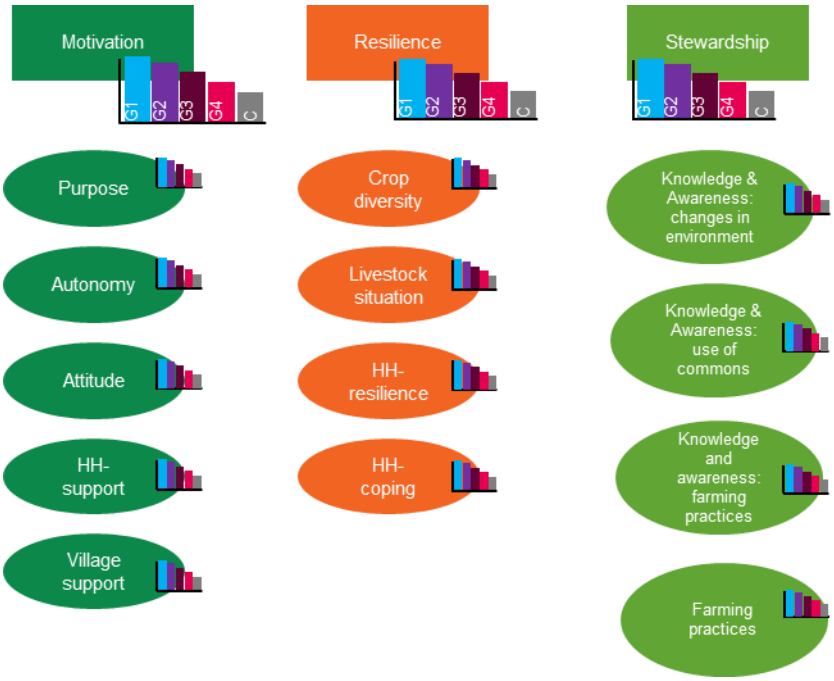


Figure 3 Schematic overview of research question 1:
Assessing effectiveness of the PIP-approach on the pillars and their sub-constructs

The outcomes motivation, stewardship, and resilience each consist of several sub-constructs that make up these main pillars. We will assess the project's impact on the pillar as a whole but also on the various sub-constructs within each pillar (research question 1.a). The expectation here is that the earlier the generation of PIP-farmer, or the longer the farmer has participated in the project, the 'higher' the scores on the outcomes (i.e. *more* motivation, *more* resilience, *more* stewardship).

2.1.2 UNPACKING THE CAUSAL CHAIN OF THE PIP APPROACH: ASSOCIATIONS BETWEEN MOTIVATION, RESILIENCE, AND STEWARDSHIP

RQ 2: To what extent do the main pillars (and their sub-constructs) of the PIP approach (motivation, resilience, and stewardship) mutually influence each other?

The second strand of research questions tries to unpack the causal chain behind the PIP-approach. Although motivation, stewardship and resilience mutually influence each other, the assumption is that motivation increases peoples' resilience and that more resilience ultimately results in increased stewardship of the land. The second research question investigates whether there are relationships between these pillars (see Figure 4). Are higher levels of motivation associated with higher levels of resilience? And do we find evidence whether high resilience corresponds with more stewardship of the land? This research question is thus about assessing the relationships (or correlations) between the three main pillars of the PIP-approach. Additionally, we also assess the associations at the level of the sub-constructs (for motivation and resilience) to see what exactly drives an increase in resilience, or an increase in stewardship. By looking at the level of sub-constructs as well, these analyses shed light on the added value of a certain sub-construct in increasing resilience and stewardship. For instance, what is the relative importance or added value of village support in increasing resilience?

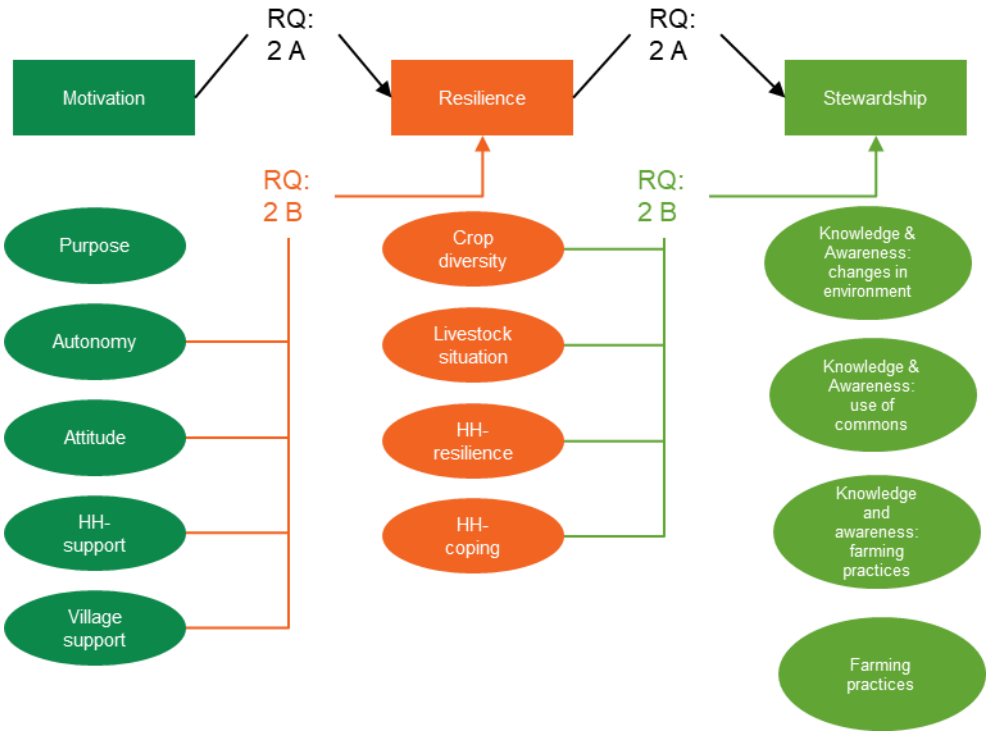


Figure 4 Schematic overview of research question 2:
Unpacking the causal chain by analysing the relationship between the main pillars (RQ 2a), and diving into more detail by analysing how sub-constructs relate to the next pillar in the causal chain.

2.2 RESEARCH DESIGN

The PABAB-consortium has already implemented the PIP-activities meaning there is no measurement available of farmers' motivation, stewardship, and resilience prior to implementing the programme. This study only thus solely uses ex-post measurements, and we can only analyse the situation after the programme has been implemented. This impact study, therefore, employs a so-called counterfactual approach using a quasi-experimental design. We use a comparison group to pinpoint the effects that can be attributed to the implementation of the PIP-activities. The comparison group serves as a reference group, and by doing so, the design mimics a true experiment. The comparison group serves as a so-called counterfactual and represents what would have happened in case the PIP-approach was not implemented.

The design is as follows. First, a mirror image (cross-section) of each generation (the factual) and a group of non-participants are sampled (the counterfactual). Each generation is matched to a similar group of farmers in the comparison group. The impact of the activities is then estimated through statistical analyses techniques comparing all generation of farmers to their respective comparison groups for all relevant outcome indicators. By hypothesizing on the pattern of differences between the various generations of PIP-farmers on the outcome indicators - i.e. the earlier the PIP-generation, the more motivation, resilience, and stewardship, and the bigger the differences with generation C (the comparison group, or counterfactual) - this design allows for an accurate assessment of the projects' impact on the pillars and their sub-constructs.

The most important limitation of this design is that it cannot establish a true causal pattern as there is no pre-project and after-project measurement available. We can only assess this at a single point in time. Assessing causality would require a true experimental set-up which is very hard to achieve given the selection process and competitions that are key to the PIP-approach. Nevertheless, the assessment of relationships among pillars, the analyses of differences between the generations of PIP-farmers using this quasi-experimental approach and counterfactual reasoning, combined with multivariate analyses does shed light on the causal pathways and mechanisms behind the PIP-approach.

2.3 METHODS AND TECHNIQUES

To assess the effects of the PIP- approach on the outcomes motivation, stewardship and resilience, we analyse to what extent these outcomes differ between a representative sample of people that participate in the project (the target group) and a comparison group. However, we know that it is very likely that the target and comparison groups are not directly comparable. They likely differ systematically on a range of characteristics. Farmers with higher socioeconomic status or relatively well-to-do farmers might be more likely to be targeted as farmers of the first generation and people that champion the PIP-approach within the broader community.

Whether farmers participate likely depends on the attributes of the household and farmers before they even join the project itself. A relatively well-to-do farmer with an innovative mindset and high aspirations and motivation is more likely to join as a first generation farmer compared to a relatively poor farmer or a farmer with a less innovative mindset who is not that motivated. The competitions that are an integral part of the approach might exacerbate such selection effects even further. Farmers who are more motivated prior to enrolling in the programme, have a learning attitude or are generally more motivated, or farmers that are relatively more well-to-do are more likely to engage in such competitions and thereby follow through with the programme. They might be better stewards of the land, more resilient farmers and have higher levels of motivation, but it is likely that this is not a result of the PIP-activities but a result of the selection process.

We thus need to correct for such selection effects to make sure we are making a fair comparison and take a comparable group of farmers that did not participate as a benchmark for those who did join to

measure the impact of the project on the outcomes of interest. In econometric terms, this means the probability of participating in the project's activities is not equal for all farmers in the community (and unknown)[1]. This probability of being treated or targeted by the project is called the propensity score. The statistical technique we use, propensity score matching, ensures that the target and comparison group is comparable based on their socio-demographic background (age, level of education, gender of the respondent, gender of the household head), the characteristics of their farm (farm type, shares of land they own and rent, size of the land they use) and face a similar agronomic context (i.e. live in the agro-ecological zones). As these selection effects differ for each generation of farmers, we also select a different comparison group (or in econometric terms: we estimate a different counterfactual) for each generation from the wider group of non-PIP farmers we have surveyed. Thus, the levels of motivation for, say generation 1 farmers are compared with a comparison group that consists of farmers with similar socio-demographics, farm characteristics, in the same agronomic context as those generation 1 farmers. Generation 2 farmers are compared with (or matched) against a subset of similar farmers in the comparison group. This means that we end up comparing four generations of farmers with four different comparison groups. The impact of the programme for each generation is then assessed by comparing each generation to their respective comparison group.

Please see annex 2.2 for a detailed description of the characteristics we use to match and find comparable farmers in the comparison group as well as an overview of the extent to which the composition of the comparison groups change before and after matching.

2.4 SAMPLING & DATA COLLECTION

In total, 962 farmers were interviewed for this study. They was drawn as a stratified multistage cluster sample. First, we divided the population into four strata, according to the four generations of PIP farmers. This stratification ensures that a sufficient number of farmers can be analysed for each generation, and allows for statistically sound comparisons between the 4 PIP-generations. In practice, this thus means that we have 'oversampled' generation 1 and generation 2. Because generation 1 and generation 2 are relatively small, each individual farmer has a higher chance of being selected as a respondent compared to the larger generations G3 and G4. This oversampling is corrected with weighting methods in the analyses, where relevant.

Based on the registration list of farmers that participated in the programme, we sampled communes in the first stage and collines in the second stage. In each colline, a fixed number of farmers in each generation was selected to be interviewed to speed up the fieldwork. The enumerators used lists of randomly selected farmers (and reserve lists in case interviews could not take place) in the selected communes and collines.

The comparison group was sampled in communes and collines adjacent, or close to the collines where PIP-farmers resided. This was done to ensure the comparability between the groups as they experience similar agronomic and market conditions as the PIP-farmers. The comparison collines were selected at random out of lists of adjacent and neighbouring communes and collines where the PIP-approach was not implemented. Enumerators were trained to conduct random walks using the widely used EPI-sampling method (or spin the bottle method) that mimics a random sample. The total number of farmers in the comparison group and thus in non-PIP collines is slightly more than half of the total sample (n=473). See figure 5 for a schematic overview of the structure of the sample and Figure 6 for an overview of the locations where interviews were conducted.

[1]Compare this to a situation where participation in the project would be determined by a coin toss (a randomized experiment). In this case, participation in the project would be solely determined by chance, not by any pre-exisiting characteristics of the people that intend to participate in the project. In this case the propensity score (the probability of being the in the target group) would be known and equal to 0.5

Project staff did not influence the choice of communes, collines, or respondents for this survey in any way, thereby limiting the possibility of cherry-picking communes, collines or respondents among whom the project was deemed more successful. Communes, collines, and respondents were all selected at random or in a process closely mimicking random sampling in the final stage of sampling for the comparison group. The multistage cluster sample further allowed to calculate the chance of each colline, commune, and respondent in the PIP-group to be included in the sample. The weighting methods used hinge on this inclusion probability so that the oversampling in the first generations is corrected for.

We have sampled in all provinces where the project was implemented except Bubanza. The security situation did not allow for data collection at the time of surveying. Shortly before the survey was fielded, the sampling was adapted (by including more people from surveyable areas, proportional to share in the population, and by providing reserve lists of communes and collines in case the security situation would deteriorate in other areas as well).

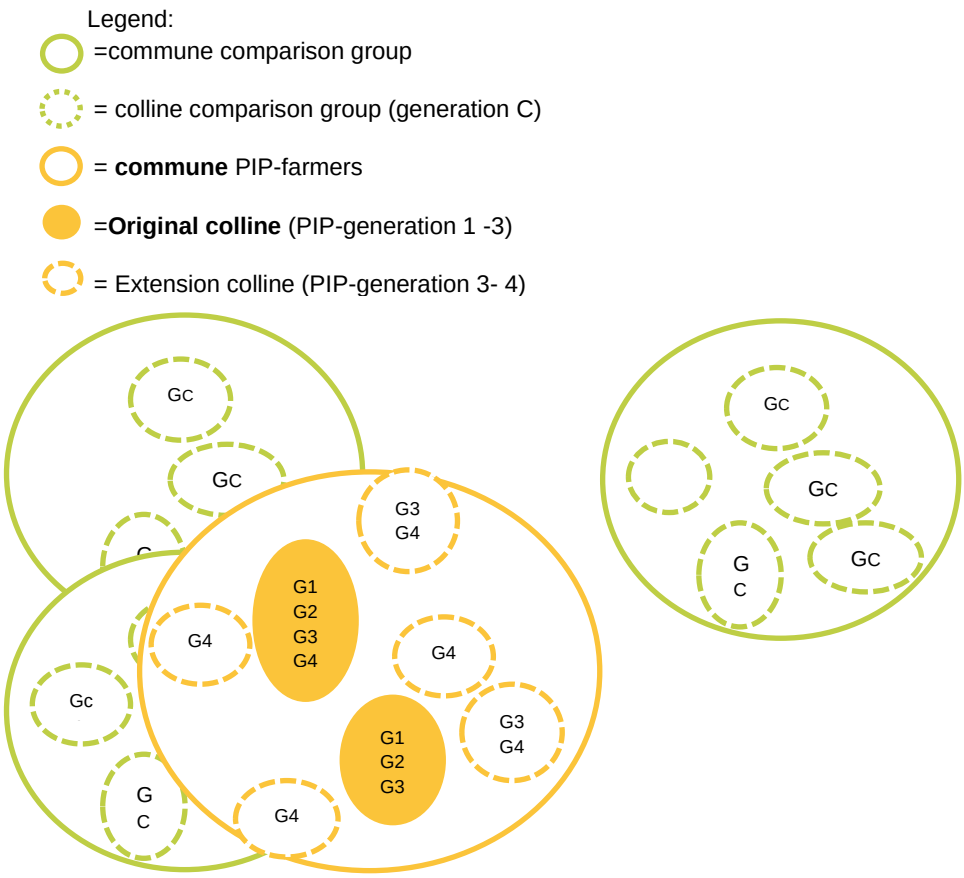


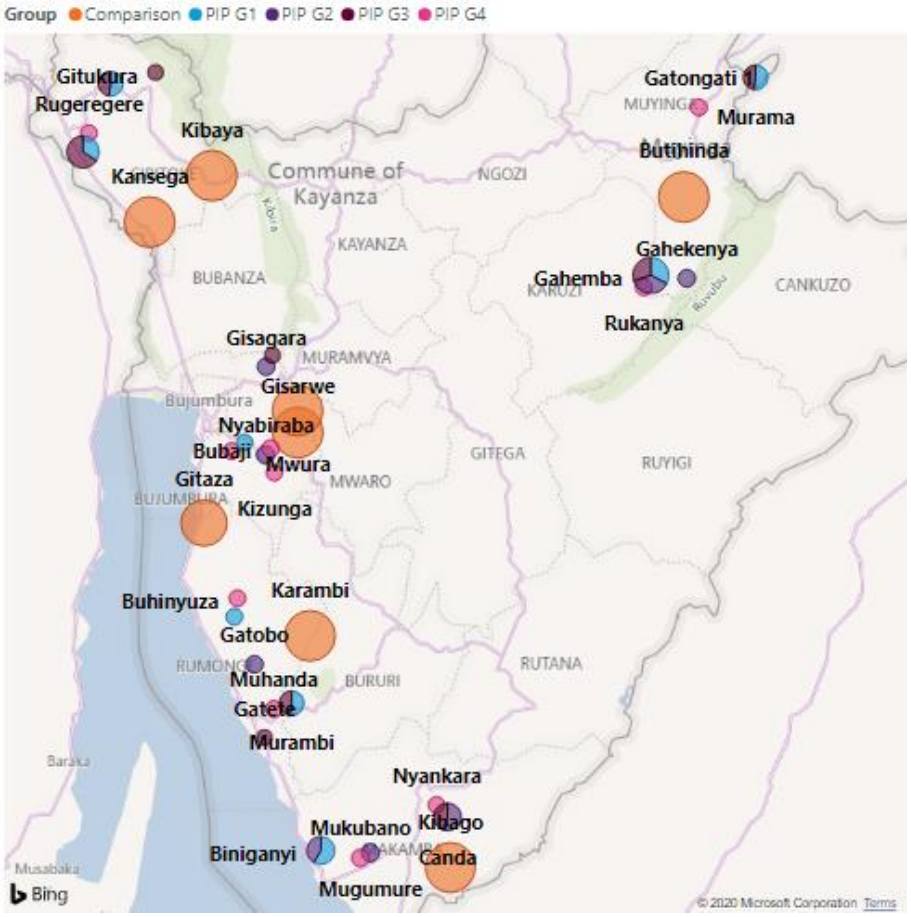
Figure 5 Schematic overview of the sampling strategy and selection of comparison communes and collines

The questionnaire for this study was developed by further building on the experience gained in previous impact measurements of the PIP-approach.

The questionnaire was developed under the auspices of Wageningen Environmental Research and builds on previous impact measurements of the PIP-approach (Kessler & van Reemst, 2018). The items in the questionnaire were extensively tested and refined using the input of SCAD-staff who trained the first generation of PIP-farmers and have been engaged with the programme throughout its

roll-out. Further enhancements were made during the enumerator training to accommodate the flow of the interview and increase the understanding of the questionnaire items among enumerators.

Figure 6: Map of sampled locations per colline and generation



Twelve interviewers were selected based on prior experience with conducting survey interviews and, or, knowledge about agricultural and agronomic practices relevant to this study and the PIP-approach. Interviewers were trained on using the survey software, interviewer techniques, sampling and understanding of the questionnaire in an enumerator training workshop in October 2019 in Bujumbura. The fieldwork was supervised by IFDC-staff. The survey was fielded between November 2019 and February 2020. Data was collected using mobile phones and tablets to aid the data entry and speed up the interview process using the software-suite Qualtrics. The training ensured that enumerators were able to work with various types of respondents and had a good understanding of the questionnaire itself and the concepts to be measured. Checks on the data collected proved that the enumerators were accurate in entering the data. There were almost no so-called response-sets, and some answer categories were randomized in order to avoid such response sets. Moreover, these checks show that interviews at the beginning of the fieldwork period lasted a bit longer (approximately 30% longer) compared to interviews conducted later in the fieldwork period confirming the expectations of the researchers. The average interview time was 40 minutes. Upon completion of the fieldwork, the data was analysed using multiple statistical software packages (STATA, Python, and R)[2].

[2] All steps in the statistical analyses process are fully reproducible. The analyses and data documentation is deposited in a github archive which can be found at https://github.com/riklinssen/papab. Source data is available on request.

3 FINDINGS

This section presents the findings of the PIP-impact study. It is structured as follows. First, we investigate who we have been working with in the project. What is the socio-demographic background of the people that we have reached? Does that differ between generations? Diving deeper into a set of generic socio-demographic, farm and household characteristics of the respondents we seek to understand better who participated in the project, whether there are differences between the generations, and which type of farmers are reached through the sequential targeting strategy of the project.

Next, we will seek to answer the first research question in section 3.2, which focuses on the effectiveness of the PIP-approach on a wide range of sub-constructs within each pillar. For each pillar, we start with a brief note on the theoretical foundations of each pillar. Subsequently, we describe how we have transferred these theoretical notions into quantitative measurements and how we have scaled and constructed each measurement before the actual results and impact of the project is described and interpreted. We end this section pertaining to research question 1 with a synthesis section where we assess where (on which pillar) the PIP-approach is most effective.

The third part of the findings section (3.3) focuses on the second research question dealing with relationships between the pillars and further unpacking the causal chain as portrayed in section 1.1. Again, we end with a synthesis section that brings the PIP-approach and the relationships between the pillars together before moving to the conclusion and recommendations in chapter 4 and 5 of this report.

3.1 WHO PARTICIPATED IN THE PIP-APPROACH?

Let us first look at who we have been working with in the PIP-project by dissecting the sample along its socio-demographics and further specifying the characteristics of the farmers, households, and farms in our sample.

3.1.1 SOCIO-DEMOGRAPHICS

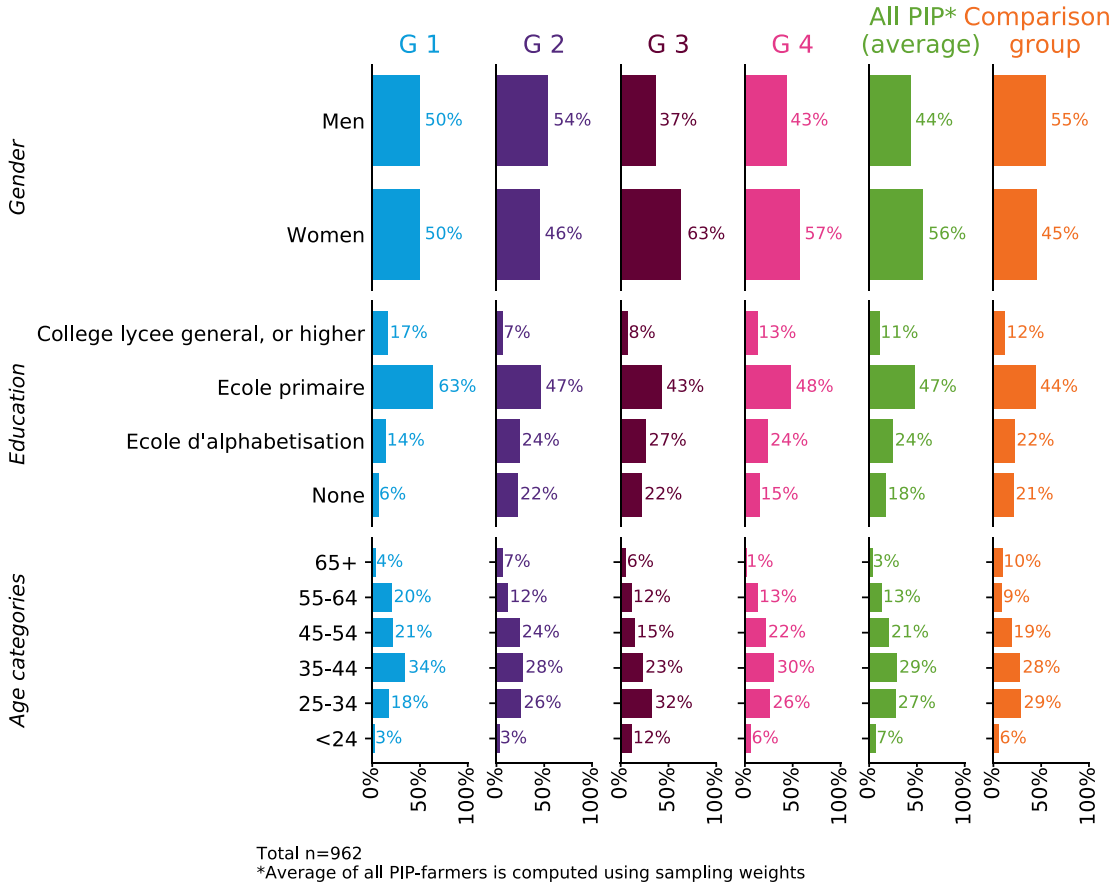


Figure 7 Socio-demographic characteristics of respondents, by PIP-generation and comparison group

Figure 7 presents a quick snapshot of the respondents of the PIP impact study. The columns represent a split by generation. The size of each bar represents the percentage of respondents within a certain category and generation. The “all PIP farmers (average)” combines all generations[3]. We have presented the comparison group (the non-PIP farmers) for reference in orange.

Female respondents are in the majority in generation 2, generation 3, and generation 4. The gender-split is equal in the first generation. Overall (in green, all PIP-farmers) we thus find a slight overrepresentation of women among the PIP-farmers surveyed. First-generation farmers are higher educated. Roughly one out of five farmers in G2 (22%) and G3 (22%) have no formal education, while in the first generation this is merely 6%. Additionally, we see that generation 1 farmers tend to be a bit older compared to the other generations of farmers.

[3] Note that the statistics for the All PIP-famers are weighted according to Generation. Because the study’s research questions requires to make comparisons between generations, we sampled an overrepresentation of G1, G2, and G 4 farmers in our sample (relative to the population) and an underrepresentation of G 4 farmers. The statistics for all PIP farmers correct for this overrepresentation of earlier generations and is representative for the population of PIP-farmers as a whole (irrespective of generation). As generation 4 is the largest group, we see that the statistics for all PIP farmers tend towards those for generation 4.

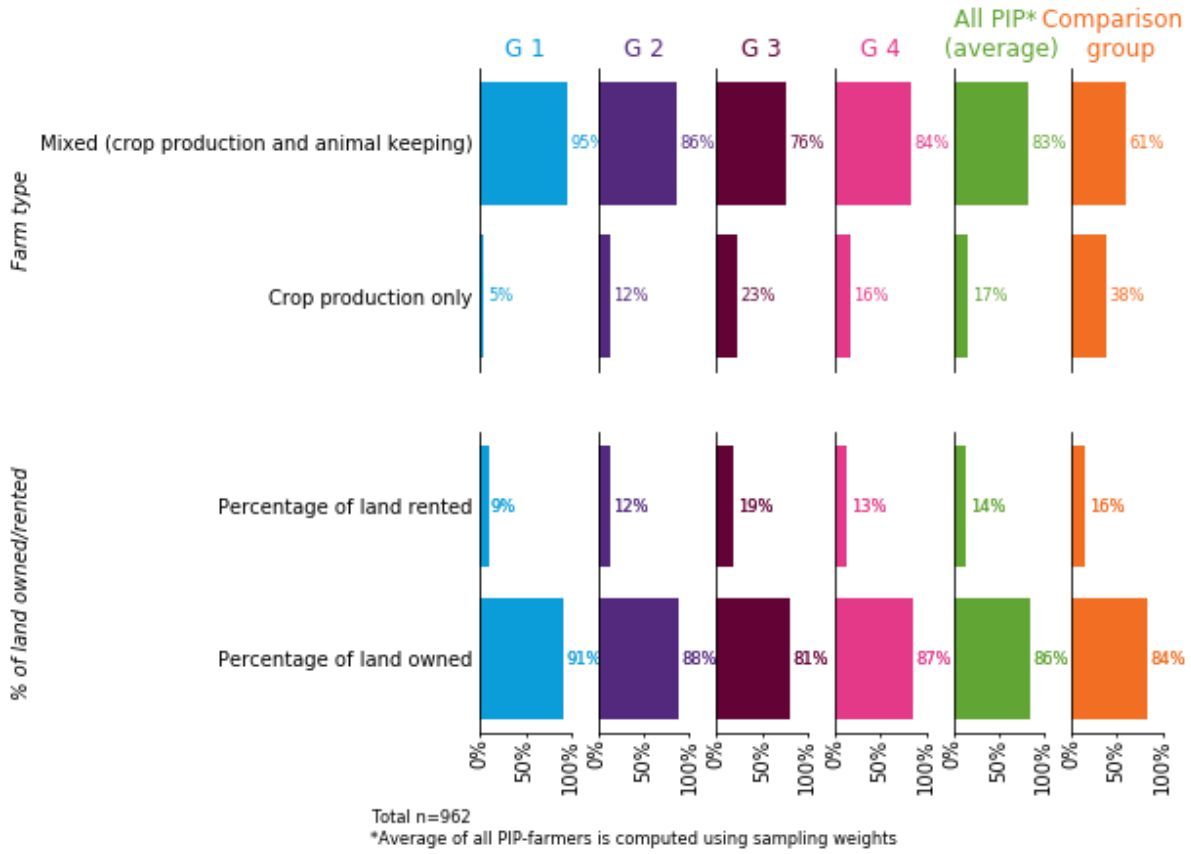


Figure 8 Farm characteristics by generation

3.1.2 FARM AND HOUSEHOLD CHARACTERISTICS

Most households have a 'mixed' farm where they combine crop production and animal keeping, shown in Figure 8. The percentage of farms that solely focuses on animal keeping is negligible in all generations and not shown in Figure 8. The share of farmers that report focusing on crop production only is 5% in generation 1, whereas this is more than three times higher in generation 4 (17%). We thus see that the percentage of farmers that solely focus on crop production is slightly higher in later generations, whereas earlier generations are more likely to hold livestock. Generation 1 is also more likely to actually own the land that they farm on, and conversely less likely to farm on rented land. From Generation 2 onwards, the percentage of land that is rented is comparable between the generations (G2, G3, G4) as well as between the PIP-farmers and the control group. First-generation farmers do not only own bigger shares of land, but the size of land they can farm on is also substantially bigger (see Figure 9). The average generation 1 farmer grows his or her crops on 2.5 hectares of land, whereas this is 1.5 hectares for generation 3.

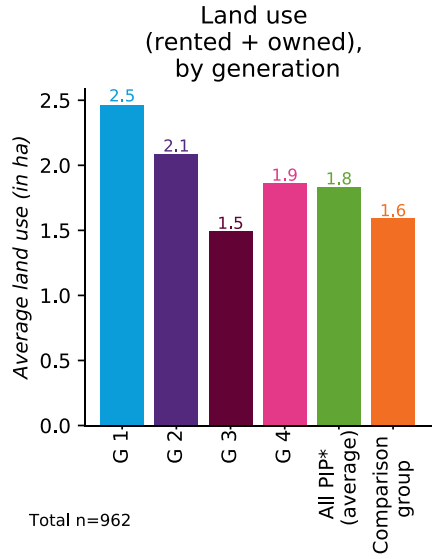


Figure 9 Land use (in hectares) by generation

All in all, G1 farmers seem to be a bit better-off and have a higher socioeconomic status. They are higher educated, more likely to own (instead of rent land) and have larger farms. The second generation follows in second place and is more similar to the first generation than other generations. The third and fourth generations are more similar to the comparison

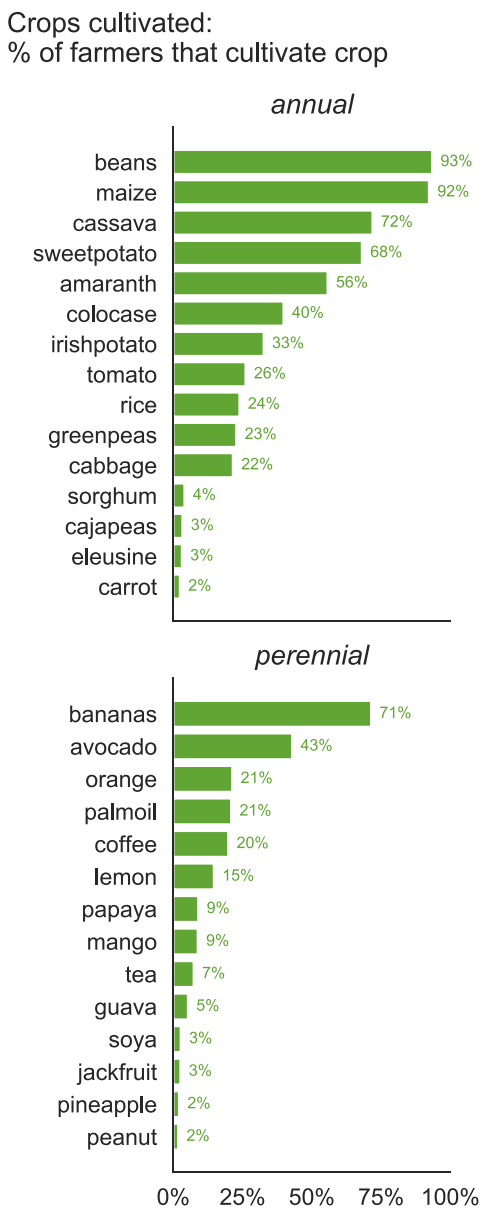


Figure 10 Annual and perennial crops grown

square in Figure 11 represents the average percentage of income gained from that particular source among all respondents surveyed. Although there is a small share of farmers that does have a non-agric income source, the vast majority of households rely on agriculture or income-generating activities that closely related to agriculture. Households with more diversified income sources, thus those less reliant on agriculture or other agric-related income-generating activities are the exception to the rule.

The focus on subsistence farming is further amplified by the large shares of produce on the group, meaning that these groups are more akin to a typical (non-PIP) Burundian farmer. The third generation of farmers is often lower on the socioeconomic ladder compared to the fourth generation. The land that they use to farm on is a bit smaller, and there are more farmers that do not hold livestock and solely focus their farm on producing crops in generation 3. Their average level of education is also lower compared to the fourth generation.

Most farmers cultivate beans, maize, banana, cassava, sweet potato, as main staple crops (See Figure 10). Crops that are typically used as cash crops are Irish Potato and perennials such as coffee, palm oil, and tea.Farmers cultivate a variety of crops, note that not all different crops are shown in Figure 10. Earlier generations tend to combine a wide variety of crops on their farms.The first generation cultivates almost 11 different crops (on average), in generation 2 this is 10 different crops whereas whereas this is between 8 and 9 different crops for generation 3, generation 4, and the comparison group (not shown).

3.1.3 INCOME AND CONSUMPTION

If we look at all respondents income sources and the diversity therein (see in Figure 11), we see a heavy focus on (subsistence) agriculture. The survey asked respondents to estimate the share of income they derive from a variety of sources. Each

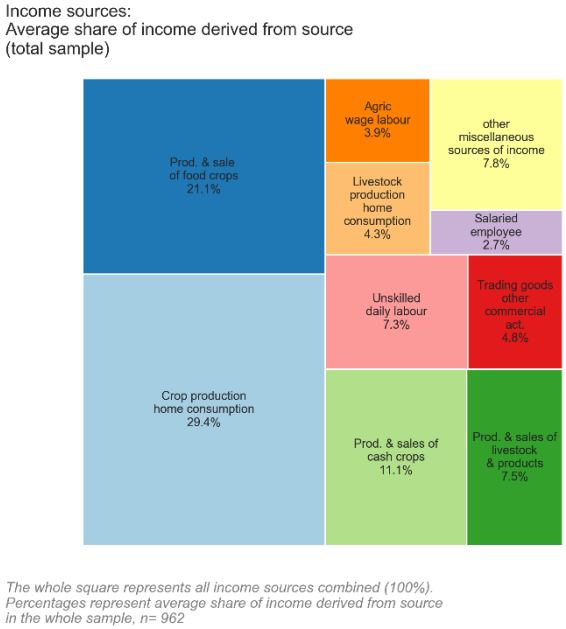


Figure 11 Average share of income derived from each income source. Average share (%) derived from source for full sample.

farm that is used for home consumption. Crop production for home consumption (29.4% light blue) and livestock production (4.3 % -light orange) already represent more than a third of total income. Other important income sources are the production and sale of food crops (which are crops that could be used for own consumption or to feed livestock but are sold at the market) which accounts for about a fifth in income. On average, the share of income derived from cash crops (crops that are grown solely for the purpose of selling them on the market) accounts for 11% of total income. Only a very small share of income is derived from non-agric sources such as skilled and unskilled (non-agric) daily labour (1.4% and 7.3%, respectively), and trading of (non-agric) goods (4.8%). Moreover, merely 2.7% of the total income is derived from working as a salaried employee.

3.1.4 PIP-PLAN COMPLETION

All respondents in the target group (generation 1 through generation 4) report to have a PIP-plan, none of the respondents in the comparison group report to have a PIP-plan, which is in line with our sampling strategy (see figure 12).

Almost one out of six farmers (17%) in the comparison group (farmers living in the vicinity of and/or neighbouring collines of those where PIP-activities were implemented) report that they 'have heard' about the PIP-approach. This indicates that some familiarity with the PIP-approach travels (likely through word of mouth) to farmers in neighbouring communes and collines as well.

The socio-economic and socio-demographic pattern we see is in line with the targeting strategy of the project. The project started with a selection of relatively well-to-do farmers that have a comparatively higher socioeconomic status (generation 1). These farmers have been directly trained by PAPAB-staff. The first generation then decides which households participate in the PIP competition and hence becomes the next generation of PIP farmers. Subsequently, the first and second generation select the third generation.

Such a targeting strategy is a double-edged sword. On the one hand, farmers that are relatively well-off and have a higher status within their community might lead by example and champion the innovative approaches they implement. On the other hand, the competitions and self-selection run the risk of the approach becoming too selective. Earlier generations of farmers are likely to select farmers more similar to themselves (thus relatively well-to-do). The competitions might then aggravate this selectiveness, which would, in turn, lead to a smaller number of farmers that enter the project at the lower end of the socio-economic ladder.

Like birds of a feather that flock together, the patterns shown above do suggest that the first generation select second-generation farmers with a similar socio-economic status. Probably because they select farmers within their own social sphere, however, we do see that the PIP-approach is able to carry through towards farmers with relatively lower socio-economic status, both in initial as well as in the extension collines. Generation 3 has the lowest socio-economic status; they are more likely to have had no education at all, have the smallest farms, and are more likely to focus on crop-production only. A large share of this group consists of farmers living in original collines (collines where the PIP-approach started with Generation 1). These farmers are often the last group of farmers in an original to join as PIP-farmers. Generation 4 is a more diverse group of farmers. They tend to be slightly higher on the socio-economic ladder compared to generation 3. Compared to Generation 1 and Generation 2, they have smaller farms they are lower educated and somewhat

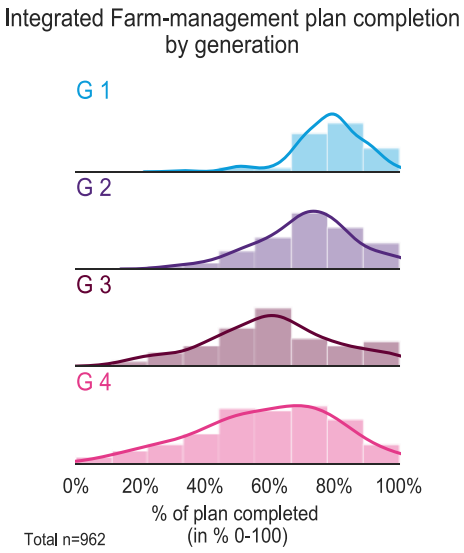


Figure 12 PIP-plan completion (in % of plan completed), by generation

more likely to focus on crop-production only. Generation 4 farmers, however, only live in extension collines, which are adjacent to the collines where the project originally started. These are thus often the first farmers that start PIP-activities in their own collines .Moreover, generation 3 and generation 4 seem to be more similar to our reference group of non-PIP farmers.

The PIP-approach with its open competitions eventually reaches about 80% of all households within a colline. The pattern that later generations are of relatively lower socio-economic status, and generally more comparable to a reference group of typical Burundian farmers within the same agro-ecological zones, demonstrates that these competitions and demonstration the PIP-approach is inclusive in later generations. These demonstrations and competitions do reach farmers at the lower end of the socio-economic ladder in later generations. These findings suggests that earlier generations indeed do lead by example.

3.2 IMPACT OF THE PIP APPROACH ON MOTIVATION, STEWARDSHIP, AND RESILIENCE

Motivation, stewardship, and resilience are multifaceted concepts, and each of them comprises of several sub-constructs or sub-components. To take into account this multifaceted nature of each pillar, we distinguish several sub-constructs for pillar motivation, stewardship, and resilience.

The operationalisation of these concepts is a process that iterates between theory and empirics. The point of departure is a theoretical understanding of what constitutes, for instance, motivation. This is operationalized into items in a questionnaire during the survey-design process. In our scaling procedures, described in the Annex, we investigate whether what we think constitutes motivation (i.e. our theoretical understanding) actually fits the data collected, by analysing the associations between questionnaire items and sub-constructs. The measures of association between farmers' responses on the questionnaire and combined initial theoretical understanding then form the basis upon which we combine items into sub-constructs and scales. By doing so, we aim to increase both the reliability and validity of our measurements of motivation, resilience, and stewardship.

3.2.1 MOTIVATION

Motivation is an inspiration or impetus to act (Ryan & Deci, 2000). Often there is a distinction made between intrinsic and extrinsic motivation. Extrinsic motivation is driven by external rewards (or punishments) whereas intrinsic motivation refers to people's internal interest and whether one finds an activity inherently enjoyable. The PIP-approach predominantly focuses on increasing farmers intrinsic motivation, thereby aiming to attain a more genuine engagement with their land, farms and the project itself/. It tries to avoid that farmers solely take actions that are driven by external rewards such as money or in-kind incentives (extrinsic motivation).

Measuring motivation

The pillar motivation consists of the five sub-constructs purpose, autonomy, attitude, household support, and village support as shown in the schematic representation below.

Attitude measures farmer's openness to innovation and whether he or she is actively seeking new farming practices to apply. Autonomy is the farmer's sense of independence and asks farmers whether they can make their own choices in life. Farmers who are more open to innovation and feel that they can make their own choices in life are more likely to be motivated. The concept of purpose is closer to the actual implementation choices farmers make to improve their livelihood. It refers to whether farmers have actively made plans and acted upon their aspirations.

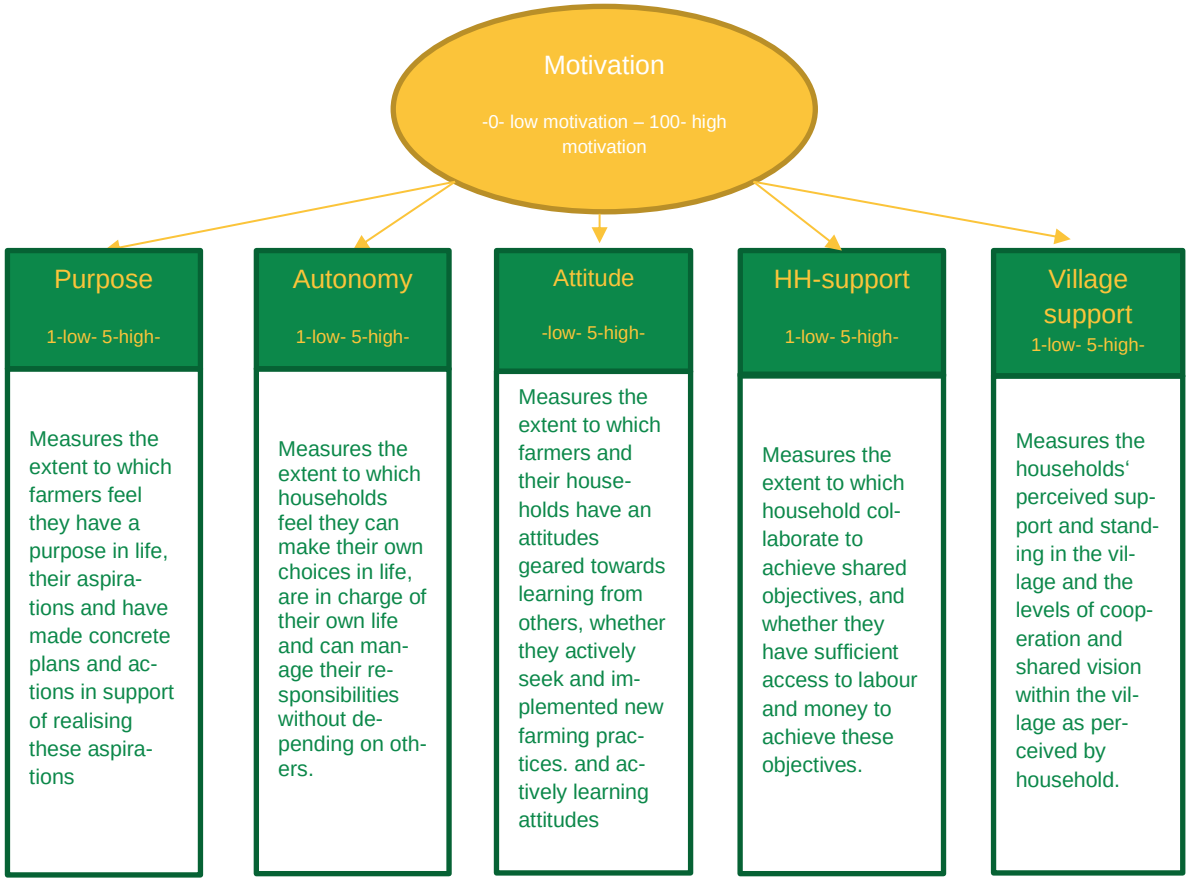


Figure 12 Measurement of motivation

Farmers do not make choices in isolation but within the context of their family and community. They thus take into account their perceptions of views and opinions of other household members and people in the broader community. Household and village support thus captures whether farmers are motivated (or not) by other household members and village members, who thus form an enabling context for being motivated to become a better farmers, and are supportive or pose barriers to the aspirations they have. Please see Annex 2 for a detailed overview of which survey questions relate to which sub-construct, the weight assigned to each survey question and how these sub-constructs are scaled.

Does the PIP-approach affect farmers motivation?

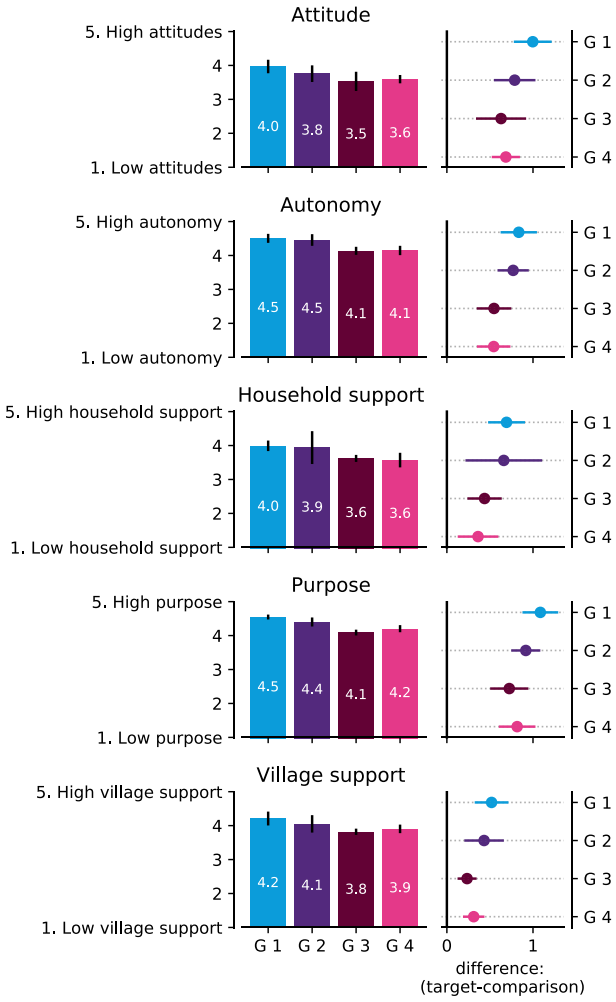
Our analyses focus on two core questions. First, is there a difference between generations of farmers' purpose, autonomy, attitude, household support and village support? Second, does the PIP-approach have an impact on farmers' purpose, autonomy, attitude, household support and village support? The results of the analyses are visualised in Figure 13. Let us first explain what is shown in these figures. The left-hand column shows the average level of attitude, autonomy, etc. for each generation

of farmers. Higher bars represent higher values of, for instance, household support. We have drawn a sample; thus, we need to take into account a margin of error around each average that we estimate[4] The vertical line in the middle of each bar represents the sampling margin or confidence interval. We assess differences in motivation between generations by comparing the bars and margins of error plotted in Figure 13. The left-hand side plots show the average value, or level, for each relevant indicator. In case the vertical bars do not overlap, there is a statistically significant difference between the generation. For instance, we see that generation 1 farmers (light blue) report significantly higher levels of household support compared to generation 4 farmers (pink).

Motivation: subconstructs, by generation

Left: values on sub-construct, by generation

Right: Impact: difference between generation and generation's comparison group



Left: Averages on subconstruct
Right: Generation- (matched) comparison (treatment effect)
Horizontal lines represent 95% confidence intervals
Non-significant differences greyed out

Figure 13 Motivation: differences between generations and impact on sub-constructs

Subsequently, we move on to the second question: assessing the impact of the project on each sub-construct. This is shown on the right-hand side of Figure 13. We measure the impact according to the counterfactual logic described in the methods and techniques section. Each generation is matched to a set of farmers that did not participate in the PIP-approach. Do note that this comparison group is not explicitly shown in Figure 13. We merely plot the difference between the comparison group and the respective generation of farmers (e.g. the score for generation 1 minus the score for generation one's comparison group). In case the value zero is not within the confidence interval shown, there is an effect of the PIP-approach' activities on that specific outcome, In case the difference is larger than zero and statistically significant, the project did have an impact on that indicator and can be considered as effective. Dots represent the difference between the generation and the comparison group in the left-hand side graphs We have plotted the 95% confidence interval around these differences (generation minus their comparison group) as well. In case the horizontal lines do not overlap with zero, the difference is bigger than zero and thus statistically significant. Or more colloquial: the project has an impact on, for instance, farmers autonomy

In Figure 13, we see that Generation 1 consistently reports the highest levels of purpose, autonomy, attitude, household support and village support followed by generation 2 and generation 4. Generation 3 farmers display the lowest value on each sub-constructs for motivation. The pattern confirms our expectation that generation 1

[4] These bars represent the 95% confidence interval. Strictly speaking we are thus 95% sure that the true value in the population for all pip or non-pip farmers lies within the range plotted, given the characteristics of this particular sample and indicator.

farmers, who are involved in implementing the PIP approach for a more extended period, display higher levels of all sub-constructs of motivation compared to other generations.

In terms of impact (visualised on the right-hand side as the difference between the generation's score and that particular generation's comparison group's average score on a sub construct), we see that the project does have an impact on all sub-constructs analysed here. Hence, participating in the PIP-approach increases the farmer's purpose, their attitude, and autonomy, as well as on household and village support they experience.

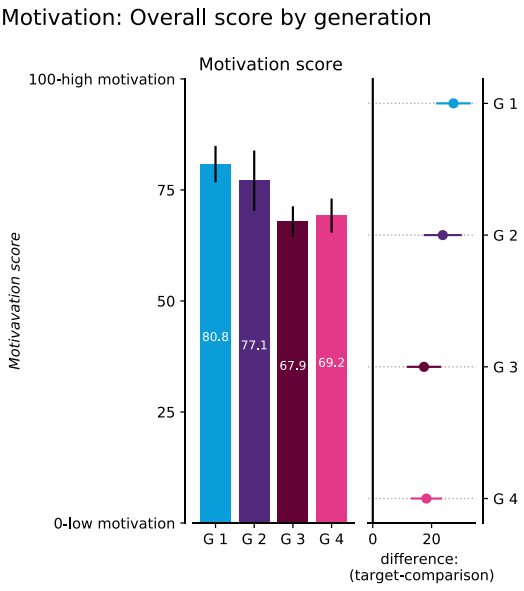
Note that all sub-constructs all have a scale ranging from 1-5; this means that, in this case, we can also compare the strength of the effects across the sub-constructs. We see that the project's impact on attitude, autonomy and purpose, is bigger compared to the impacts village support (and to a lesser extent household support). The relatively small effect on village support might be because the support of the village is not as much in a farmers sphere of control compared to their own sense of purpose or autonomy.

Finally, we find a similar pattern in *levels* of autonomy, attitude, etc. as we do in the *impact* of the project on autonomy, attitude, etc. Where older generations (Generation 1 and Generation 2) display higher levels (i.e. the bars for these generations are consistently higher than for the others) of these sub-constructs, we also see that that the projects' *impact* on these sub-constructs tends to be stronger for earlier generations (G1) compared to younger generations (G4).

Next, we bring the multifaceted concept of motivation together and move to the pillar level. Here we create an overall motivation score that includes all relevant sub-constructs purpose, autonomy, household support, attitude, and village support. This motivation score is constructed as a weighted average of all sub-constructs presented in Figure 12 and ranges between 0, which means the lowest motivation and 100, representing the highest level of motivation[5].

These findings are presented in Figure 14. The pattern in levels of motivation is equivalent to that seen earlier when separately assessing the various sub-constructs. The first generation scores the highest followed by the second generation, and then down towards generation 4 and generation 3.

If we look at the impact of the project on motivation (right-hand visual), again we see a familiar pattern. The project has a big impact on motivation. The impact of the project seems to be bigger for earlier generations compared to later generations. Generation 1 farmers report an average motivation score of 80.8, against a score of 52.6 (not shown) among a comparable group of farmers that did not participate in the project (the comparison group for generation one). The motivation score is thus about 1.5 times higher for farmers in generation 1 compared to those who do not participate in the PIP-approach. Generation 4 reports an average motivation score of 69.2 against a score of 50.95 (not shown) for their comparison group. Thus, even the fourth generation



Left: Averages motivation score
Right: differences Generation- (matched) comparison (treatment effect)
Horizontal/vertical lines represent 95% confidence intervals

Figure 14 Overall motivation score: differences between generations and impact

[5] Please see annex 2 for a more detailed description of how this overall motivation score is constructed.

of farmers in the extension collines, which has been participating in the programme for a shorter period, and where the bulk of participants is only halfway through completion of their PIP-plan (see Figure 12), reports levels of motivation that are more than 1.3 times higher as those who do not participate in the PIP approach.

All in all, these results show that the project's activities are successful in increasing farmer's motivation, for all generations of farmers. The change in motivation as a result of the project is also bigger for earlier generations. Moreover, we see stronger impacts of the programme on the elements of motivation that are within the sphere of influence of the farmer (attitude, autonomy, and purpose). The effect of the project on perceived village support, which is arguably outside the farmer's sphere of influence, is somewhat weaker but still statistically significant.

3.2.2 RESILIENCE

Resilience is defined as the ability of a system to return to its initial state after a shock or perturbation (Holling, 1973). Resilience in the PIP-approach refers to the ability of farmers to recover from - or adapt to - sudden changes in the environment and exogenous shocks. It has both a physical and behavioural dimension and a dimension focusing on adaptability of households to unexpected shocks that is more attitudinal in nature.

The physical dimension refers to whether farmers diversify the crops they grow and livestock they hold. A diversified farm that produces a broader range of crops and livestock products and by extension sells a wider range of products on the market is less likely to suffer from the adverse effects of exogenous shocks.

Adaptability refers to the household's skills, knowledge and levels of collaboration that determine whether farmers can minimise the impacts of events that adversely affect their livelihood. The project, therefore, focuses on increasing the households adaptive capacities and improving intra-household household decision making when farmers formulate their PIP-plan.

Measuring Resilience

The physical dimension is measured as a farm's crop and livestock diversity by counting the number of different crops they grow, livestock they hold, and crops or products they sell at the market. For livestock, we have also included the farmers' assessment of fodder sufficiency for the livestock they hold. Figure 15 gives a schematic overview of the measure of resilience.

The adaptability dimension is operationalized in two sub-constructs; household resilience and coping ability. First, household resilience includes questionnaire items that measure the way how intra-household decision making takes place. Furthermore, it adds a subjective appraisal of whether households have sufficient access to skills and resources to handle unexpected events, and a subjective assessment of the health situation of household members. Second, coping ability is measured by asking respondents how they dealt with the last shock they experienced[6] and whether they have learned from it and if they would do anything different right now. Please refer to Annex 2.2 for a detailed overview of which survey questions are posed to respondents and the relative weight assigned to each item[7].

[6] Note all hh- experienced a shock at least one shock out of 12 shocks mentioned, see the questionnaire for an overview shocks mentioned.

[7] Most conceptualisations of resilience emphasize the importance of income diversity for resilience. A household that relies on a more diversified range of income-sources would indeed cope better with adverse events because losses incurred from for instance a lower yield during the harvesting season could be offset by incomes from a different, non-agric source. There are only a very small number of households in the sample that have income sources other than those directly related to agriculture. Al-

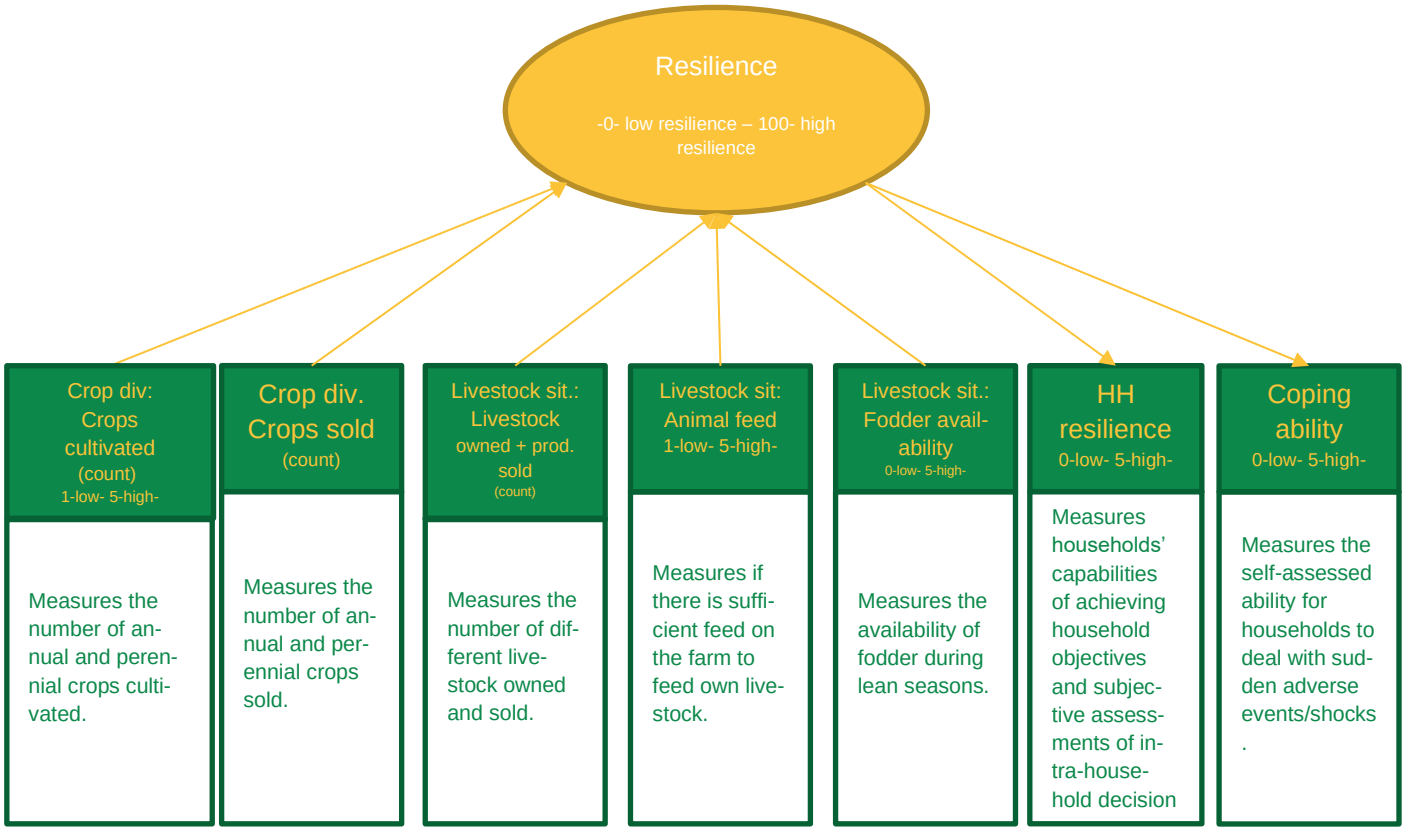


Figure 15 Measurement of resilience

Does the PIP-approach affect farmer's resilience?

The impacts of the project's approach on crop-diversity are presented in Figure 16. The presentation and interpretation of these visuals are analogous to those explained in the motivation section. Note that effects (differences between target and comparison groups) that are not statistically significant are greyed out.

Let us first look at the physical dimension of resilience that is livestock and crop diversity. We see the familiar pattern that the first generation generally reports the highest levels of diversity, followed by the second generation and the third and fourth generation. However, we find that PIP-farmers do not cultivate a wider variety of crops compared to non-PIP farmers. With the exception of generation 1, PIP-farmers neither sell a wider variety of crops at the market compared to non-PIP farmers. We thus find that although generation 1 sells a wider variety of products on the market, the PIP-approach does not directly lead to an increase in crop diversity.

Generation 1 and generation 2 own a wider variety of livestock compared to their respective comparison groups and report to have sufficient fodder resources available on their farms (see figure 17). Hence, the PIP approach is only associated with a wider variety of livestock and the ability to properly manage livestock for generation 1 and generation 2.

---

most all (98%) respondents are farmers, and the median share of income derived from agric-related sources is 100%. Because households with a diverse income are exceptional cases in this particular sample we have excluded income diversity from our conceptualization and measurement.

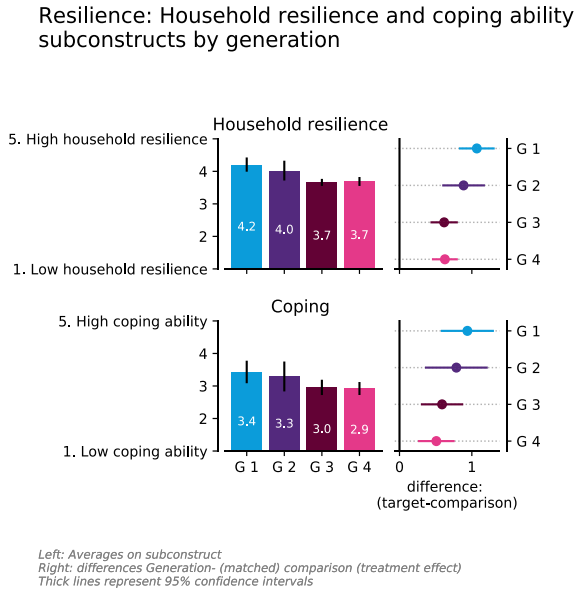


Figure 18 Resilience: household resilience and coping ability

To increase the number of different types of livestock held; one needs to have livestock in the first place. Recall from section 3.1, that the share of farmers that do not hold any livestock at all and solely focus on crop production is highest in generation 3 and generation 4. Transitioning to a mixed farm first, and then increasing the diversity of livestock held take substantial investments in terms of both time and money. The farmers in Generation 1 and generation 2 are already more likely to have moved towards a mixed farm and have had more time to diversify the types of livestock that they hold (as they have also been participating in the project for a longer time).

Farmers in generation 3 and generation 4 that do hold livestock do not differ significantly from their comparison groups.

We do find impacts of the PIP-approach on the fodder resources farmers report to have available during the lean season for all generations except generation 3. The availability of fodder resources during the lean season can be interpreted as a stricter litmus test for a farms resilience compared to the other indicators; it captures the extent to which resources are available under duress rather than the availability of resources under normal circumstances. With the exception of generation 3, all generations report having more resources available compared to their non-PIP counterparts.

Let us now move to the adaptability dimension of resilience that consists of the households resilience and their coping ability. The results are shown in Figure 18. We see positive impacts of the project on both household resilience and a households coping ability. Thus, PIP-farmers have organised their decision making around farming inputs and which crops to grow better compared to non-PIP farmers. They have more knowledge about integrated farm management and can depend on more resources in case problems arise. Similarly, we find that PIP-farmers state that they are better able to cope with events potentially detrimental to their livelihoods compared to non-PIP farmers. The pattern in differences between generations is all too familiar by now; Generation 1 reports the highest levels, followed by generation 2, down to later generations 3 and 4.

All in all, the effectiveness of the PIP-approach on the physical dimension of resilience (read: crop and livestock diversity) is rather mixed. We only find more crop and livestock diversification in generation 1 and generation 2. However, the project is effective in increasing farmers adaptability to shocks. PIP-farmers unambiguously report higher levels of household resilience and better coping ability with exogenous shocks than non-PIP farmers.

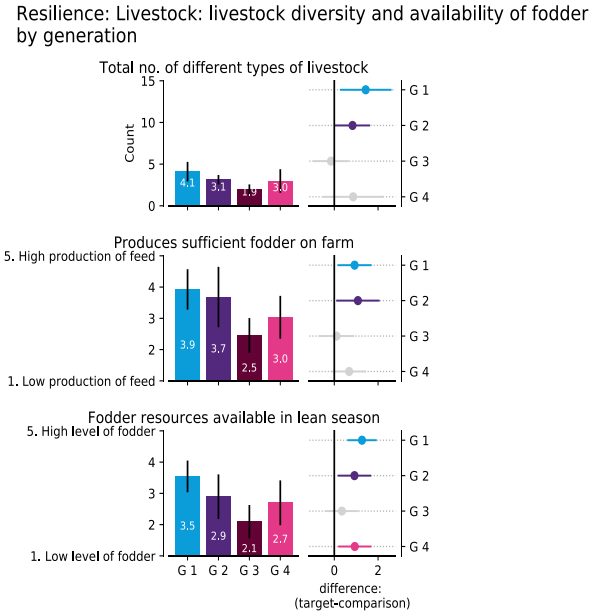
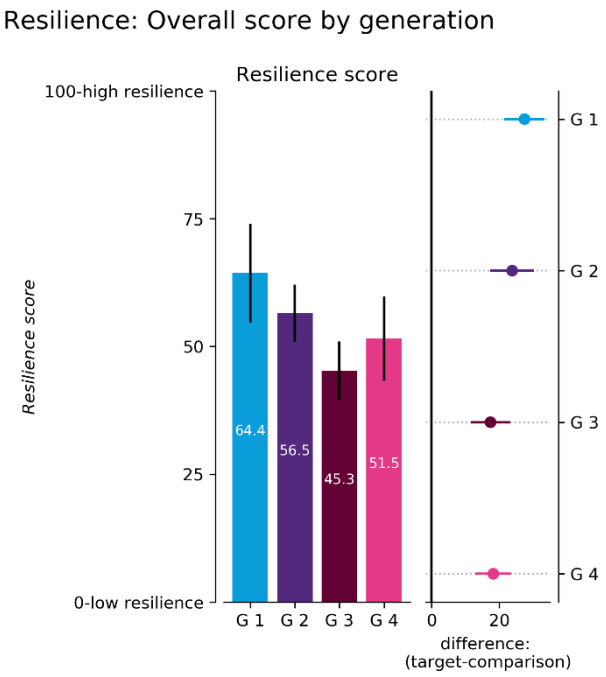


Figure 17 Resilience: impacts on livestock diversity

The analyses demonstrate that the PIP-approach is indeed effective on the adaptation component of resilience, but the effect on changing the physical component of resilience is limited. An explanation might lie in the fact that adaptation is more related to the stance people take towards mitigation of adverse events. The physical dimension requires that such a stance or attitude materialises in actual behaviours and adaptations that require substantial investments in livestock and coming to grips with growing a wider variety of crops. Such adjustments likely need a longer time to achieve, and in the case of livestock diversification these changes require substantial investments. This also explains that the PIP-approach seems to be effectively increasing the physical dimension for generation 1, and generation 2 only. These farmers have participated the longest, and most of them have implemented their plans almost to the fullest (see section 3.2.1).



Left: Averages resilience score
Right: differences Generation- (matched) comparison (treatment effect)
Horizontal/vertical lines represent 95% confidence intervals

Figure 19 Resilience: impact on resilience score

Figure 19 shows the levels and impact of the project on resilience as a whole, disaggregated by generation. The overall resilience score is a weighted average of the sub-constructs described previously and ranges between zero (low resilience) and hundred (high resilience). See Appendix 2 to learn more about the construction of this index. The project effectively increases the resilience of those who participate. The first generation reports not only the highest levels of resilience, but they also show the biggest increase in resilience as a result of the project's activities compared to non-participants. The change in resilience is smallest for the third and fourth generation of participants.

3.2.3 STEWARDSHIP

Stewardship refers to our responsibility to manage and protect the land and its natural heritage (Brown & Mitchell, 1998). Stewardship goes beyond the mere application of a set of sustainable farming practice and management of land and farm. Management would imply a focus on efficiency and profitability. Instead, stewardship comes with the mindset of applying such methods for a benevolent purpose or even motivated by feelings of moral duty. The PIP-approach fosters stewardship of the land through creating awareness of changes in the natural resources and the environment of the farm, the prevention of pollution, soil degradation, and erosion. PIP-farmers enhance their capacity to be good stewards of the land in workshops and training where they share their views and knowledge with other farmers. Moreover, these workshops and pieces of training present tools and best practices that farmers can apply to manage their farm, conserve land, and sustainably use the commons.

Measuring stewardship

Our measurement of stewardship is based on several measures capturing farmers knowledge and awareness as well as a behavioural component that captures whether farmers actually implement a wide variety of farming practices. We measure peoples knowledge and awareness about changes in the environment, their knowledge about conserving the commons. Moreover, the survey captured the respondents' knowledge on when, how, and why to implement farming practices.

The questionnaire included questions on a set of farming practices. The first are land management practices, which are non-permanent changes to plots themselves that are implemented annually or with each growing season of a particular crop, such as ploughing on the contour line, staggered row planting of crops, the use of mulches, and planting cover crops to combat erosion. Second, we measure whether farmers apply a set of physical practices. These are usually more permanent changes to land and farm, such as digging contour lines (i.e. trenches on the contour lines to capture runoff water) and continuous ridges (also on the contour lines), or the implementation of some forms of gully control. Third, and in line with one of the core focal points of the PAPAB-programme, the survey inquired about the type and combinations of fertilizer farmers use (manure, compost, chemical fertilizer) or whether respondents use a combination of these fertilizer practices. Finally, we inquired whether farmers used a form of crop-ration on most or all of their plots. A schematic overview of the stewardship pillar is presented in Figure 20. Analogue to the previous structure we first investigate the PIP-approach impact on these sub-constructs separately (for each form of knowledge and each farming practice) and finally take an all-encompassing look by studying whether PIP-farmers display higher values on the overall stewardship pillar compared to non-PIP farmers.

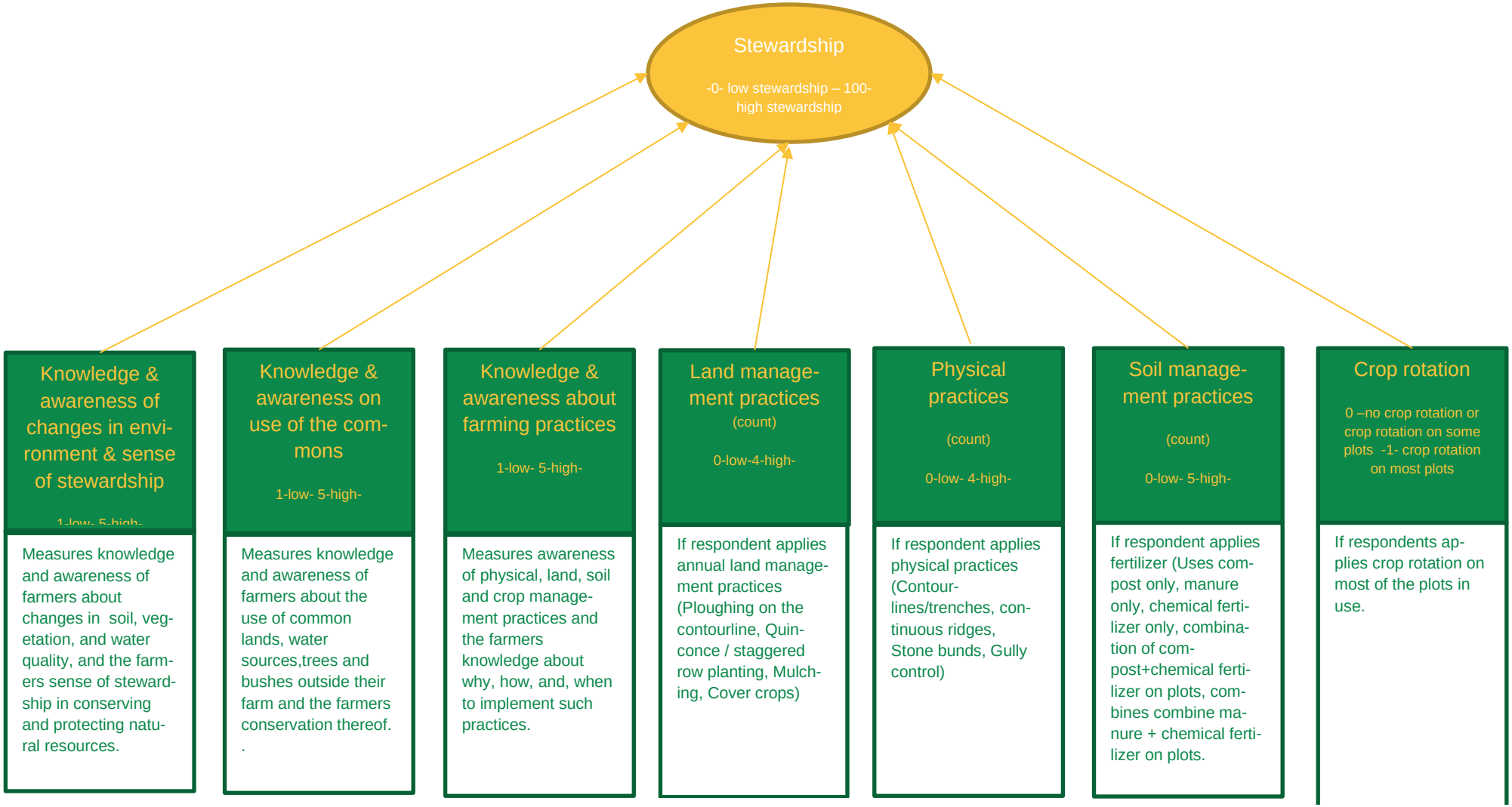


Figure 20 Stewardship: measurement

Does the PIP approach affect farmers' stewardship?

Let us first assess the knowledge and awareness elements of stewardship, presented in Figure 21. We find that PIP-farmers are more aware of changes in land, soil and their natural environment, and that they have more knowledge about how to conserve and protect natural resources (sense of stewardship), compared to non-PIP farmers (right-hand side, Figure 21). Similarly, PIP-farmers have more knowledge about the use of the commons. Additionally, PIP-farmers have more knowhow about how and why to implement various farming practices.

The levels of knowledge and awareness are highest in the first generation and lowest in the third and fourth generations, which is in line with the general pattern we have found for indicators resilience and motivation.

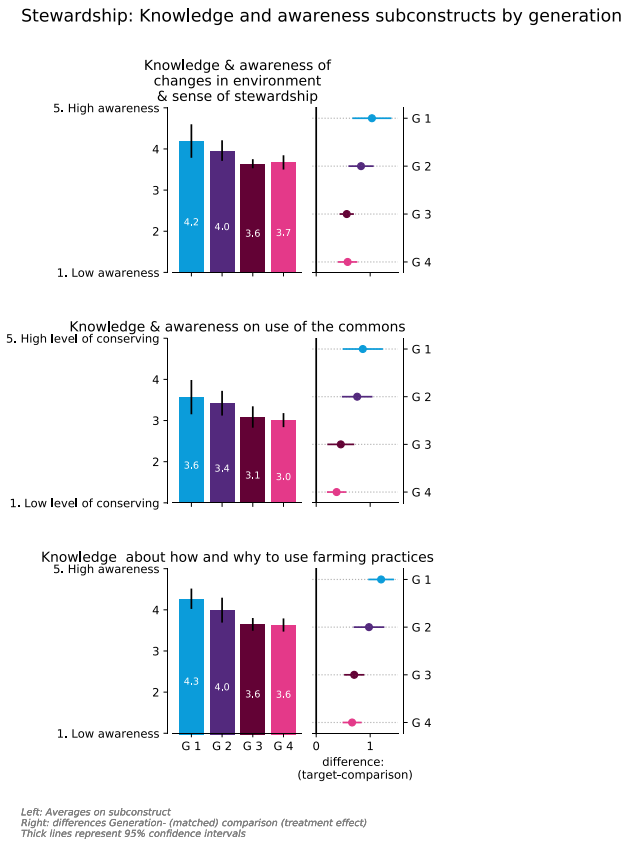


Figure 21 Stewardship: impacts on knowledge and awareness of changes in the land & soil, sense of stewardship, and knowledge and awareness of farming practices.

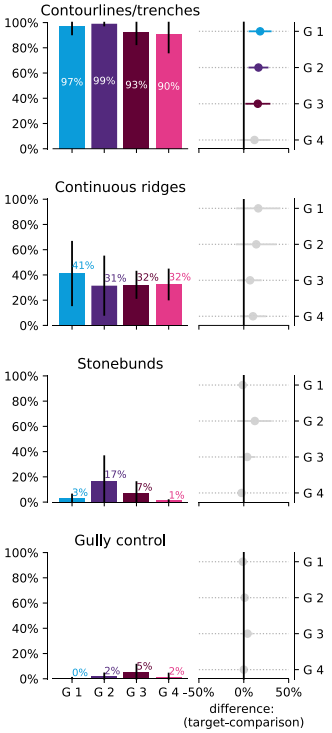
We find positive impacts across the board, meaning that those farmers who participate in the PIP-approach are more aware of changes in their environment, and especially more knowledgeable about a variety of farming practices. Additionally, PIP-farmers have a better awareness of changes in the natural environment that might affect them, and they have more knowledge on the use of the natural resources such as water (use of the commons). The approach thus affects all subconstructs that speak to farmers knowledge and awareness positively. The impact of the PIP-approach is strongest on increasing the knowledge regarding how and why to implement various farming practices; the difference between the target and comparison group (right-hand side) is the largest for all generations for this particular indicator.

By looking at practices farmers actually apply (instead of the knowledge farmers have about these practices), we turn to behavioral aspects of stewardship. This is shown in Figure 22, Figure 23, and Figure 24, which present a wide assortment of farming practices. In case differences between PIP and non-PIP farmers are not statistically significant meaning that we do not find an impact of the project on these practices, the elements in the figure are greyed out.

For the physical practices (Figure 22), we see that a larger share of PIP-farmers applies contour lines compared to non-PIP farmers. Virtually all PIP-farmers in the first and second generation use this practice. Although PIP-farmers are more likely to apply contourlines than non-PIP farmers, the application of trenches and contourlines is also rather commonplace among non-PIP farmers. More than four out of five non-PIP farmers also apply this practice. This is not shown in the graph but can be derived from the fact that in Generation 4, a share of 90% of farmers applies the practice. Whereas this share among the fourth generation of farmers is very high (90%), this share is similar among their non-PIP counterparts, evidenced by the absence of a statistically significant difference between Generation 4 and their comparison group. Hence, we do find that PIP-farmers are slightly more likely to implement countourlines than non-PIP farmers (except for generation 4), but the practice itself is almost universally applied among all farmers. This also begs the question whether there is more growth

possible in the use of countourlines and trenches. The application of other physical practices is relatively rare, especially stonebunds and gully control, and to a lesser extent the use of continuous ridges. Stonebunds, gully control, and continuous ridges are rare among PIP-farmers, and equally rare among Non-PIP farmers. There is no impact of the PIP-approach on the implementation of these physical practices.

Stewardship: Physical practices:
% of people that applies practice (left)
difference between generation and comparison group (right)



Left: % of people that applies practice
Right: differences Generation- (matched) comparison (treatment effect)
Thick lines represent 95% confidence intervals
Differences that are not statistically significant (p<0.05) greyed out

Figure 22 Stewardship: Impact on physical practices

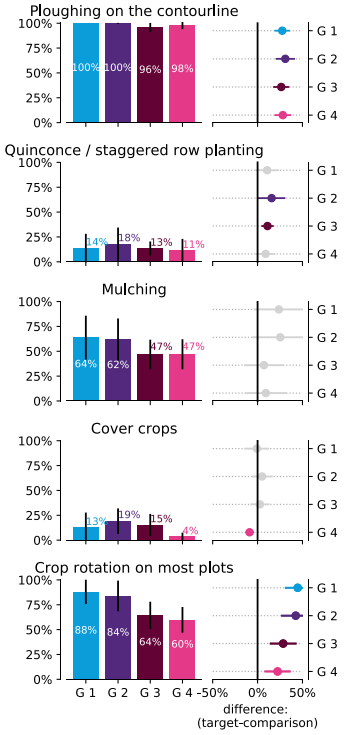
The final set of practices we look at is the use of fertilizer, presented in Figure 24. The expectation here is that PIP-farmers more consciously choose how to enrich the soil and seek to combine a variety of fertilizers to most of their plots of land. Thus we expect PIP-farmers to combine chemical fertilizer with compost, instead of (only) using manure chemical fertilizer, or compost on their land solely.

This is also exactly what we find. PIP-farmers choose to combine different types of fertilizer more often, whereas non-PIP farmers are more likely to rely on a single source of fertilizer. For Generation 3 and 4, we find a bit more nuanced picture (bottom two graphs in Figure 23, the left-hand side, where some differences are greyed out). Generation 4 is more likely to use manure compared to their comparison group, whereas Generation 3 is more likely to use compost. Recall from section 3.2.1. that almost a quarter of the farmers in Generation 3 does not own livestock, whereas this only 16% in Generation 4. Thus, it is likely that farmers in Generation 3 choose for the combination of

Land management practices and crop rotation are visualized in Figure 23. We find that PIP-farmers are more likely to plough on the contourline compared to non-PIP farmers (top of Figure 23). The same holds for crop rotation (bottom). The difference between the share of PIP-and non-PIP-farmers that ploughs on the contourline, and rotates their crops on most of their plots is very large. These are two of the practices where PIP-farmers really distinguish themselves from non-PIP-farmers.

On the contrary, staggered row planting of crops, using mulches, and planting cover crops is not more common among PIP-farmers compared to non-PIP farmers. We do find a significant difference for Generation 2 and Generation 3 in the application of staggered row planting, but this is only barely significant, and the difference between PIP and non-PIP farmers is very small.

Stewardship: Land mngnt, practices & crop rotation:
% of people that applies practice (left)
difference between generation and comparison group (right)



Left: % of people that applies practice
Right: differences Generation- (matched) comparison (treatment effect)
Thick lines represent 95% confidence intervals
Differences that are not statistically significant (p<0.05) greyed out

Figure 23 Stewardship: impact on land management practices & crop rotation

compost and chemical fertilizer, instead of manure and chemical fertilizer, because a larger share of Generation 3 does not own any livestock.

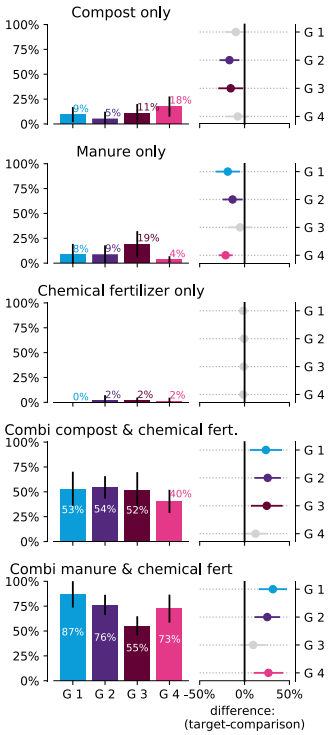
All practices taken together, we see strong impacts of the PIP-approach on soil management practices. PIP-farmers use a wider variety of fertilizers on their farms. The evidence regarding physical practices is mixed. We do not find an impact on applying most physical practices, and for those where we do find a small impact (contourlines) are almost universally applied among non-PIP farmers as well. Crop rotation is a practice where PIP-farmers really distinguish themselves from non-PIP farmers. Regarding land management practices, we only find that PIP-farmers plough along their contourlines more often compared to non-PIP farmers. Also note that a lot of the practices taken into account here are relatively rare across the board (both for PIP and non-PIP farmers), especially the physical practices.

Through analysing each stewardship sub-construct separately we have seen that PIP-farmers do have more knowledge about a variety of farming practices, the use of fertilizer and are more aware of changes in soil, vegetation and more generally their land. However, the knowledge they have about a wider variety of farming practices does not unambiguously translate into the application of a markedly different set of farming practices. Although PIP-farmers, use a wider variety of fertilizers and implement crop rotation on most of their plots, they do not differ from non-PIP farmers on most other practices taken into consideration here.

Let us now turn to the pillar stewardship as a whole and provide a broader, overarching look by analysing the overall stewardship score. This is presented in Figure 25. Figure 25 shows that all generations of PIP-farmers score higher on the overall stewardship pillar compared to non-PIP farmers. The pattern is again equivalent to that seen earlier for the pillars motivation and resilience. Generation 1 leads and Generation 2, 3, and 4 follow.

All sub-constructs considered, we thus find PIP-farmers are better stewards of their land compared to non-PIP farmers. The analyses thus lead us to conclude that the PIP-approach is also effective in increasing farmers stewardship. The effectiveness of the approach on farmers stewardship is, however, mostly driven by the PIP farmers' knowledge

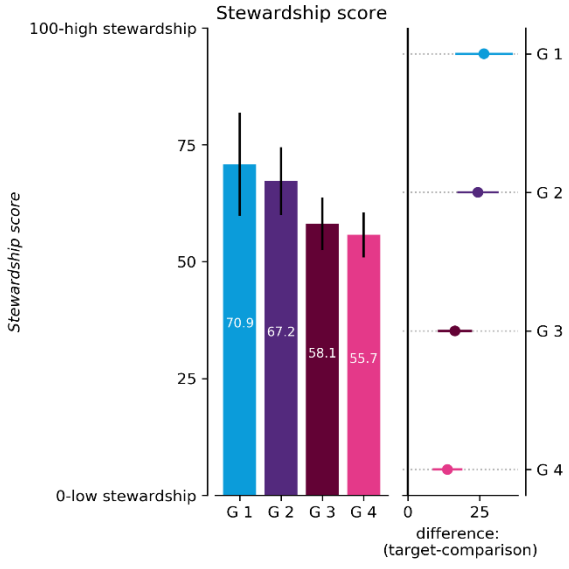
Stewardship: Soil management practices:
% of people that applies practice (left)
difference between generation and comparison group (right)



Left: % of people that applies practice
Right: differences Generation- (matched) comparison (treatment effect)
Thick lines represent 95% confidence intervals
Differences that are not statistically significant (p<0.05) greyed out

Figure 24 Stewardship: Impact on soil management practices (fertilizer use)

Stewardship: Overall score by generation



Left: Averages stewardship score
Right: differences Generation- (matched) comparison (treatment effect)
Horizontal/vertical lines represent 95% confidence intervals

Figure 25 Stewardship: impact on stewardship score

and awareness[8] about stewardship in general and knowledge about farming practices. PIP-farmers seem to be more conscious that they should be good stewards of the land, they are more likely to implement a crop-rotation scheme but do not implement a vastly different set of farming practices compared to non-PIP-farmers.

3.2.4 SYNTHESIS

Where do we find the strongest effects of participating in the PIP-approach?

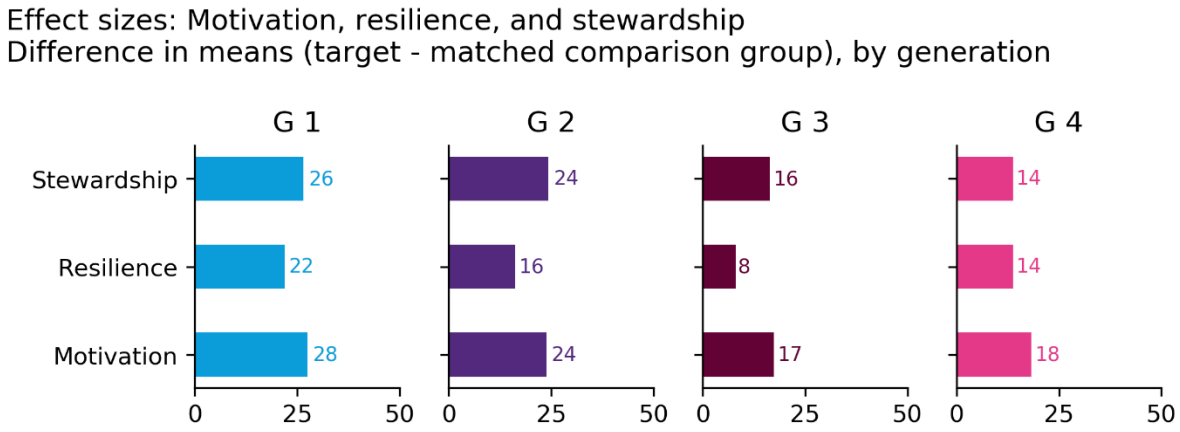


Figure 26 Effect sizes by pillar: The PIP-approach has the strongest effect on people's motivation, followed by stewardship, and finally resilience

The sections above took a deep dive into each pillar. We were seeking to understand the effectiveness of the PIP-approach on each subcomponent *within* a pillar. In this section, we will compare the efficacy of the PIP approach *across* pillars. Thus, investigating whether the PIP-approach is more effective in, for example, changing stewardship of the land than motivation.

In Figure 26, we have plotted the three pillars (stewardship, motivation, and resilience on the vertical axis) and the so-called effect size on the horizontal axis. The effect size is represented by the bars where each bar represents the difference in outcome (e.g. motivation) between farmers that participate in the PIP approach and those that do not. Thus, for instance, on the top left the score on stewardship is 26 points higher for the Generation 1 farmers, compared to a similar group of farmers that did not participate in the PIP-approach (generation one's comparison group). Furthermore, all scores for motivation, stewardship, and resilience range between 0 and 100 meaning we thus compare the relative strength of the effect of the PIP-approach on each pillar by comparing the length of the bars to each other. Vertically we can compare effect sizes across pillars, and horizontally we can compare effect sizes across pillars.

As mentioned before, we find the biggest impacts of the PIP-approach on motivation, stewardship, and resilience (comparing horizontally in Figure 26) for those that have been participating in the project for the longest (generation 1) followed by generation 2.

The programme is most effective in increasing farmers motivation, for all generations. The effects of the programme are a bit smaller on the resilience and stewardship of the land. The motivation for a farmer in generation 1 is 1.5 times higher (not shown) compared to non-participants. Although in later generations the effect is somewhat smaller, farmers in generation 4 (those that have been participating for the shortest amount of time) still report that their levels of motivation are almost 1.4 times higher than those who do not participate. Relative to the changes brought about by the project on the other pillars; the project is least effective in increasing farmers resilience. This is mainly due to the

[8] Also note that the weights for the knowledge and awareness sub-constructs is relatively high compared to the elements of practices.

comparatively small effect of the PIP-approach on resilience for Generation 3. Recall from section 3.2.1 that the third generation of farmers is also the poorest group of farmers, which indeed makes major investments that increase resilience such as investments in (more diverse) livestock a bit harder for this group. Still, there is no reason to despair; even though this is the smallest effect of the PIP-approach we observed, the resilience of the generation 3 farmers is still 1.2 times higher compared to those who do not participate.

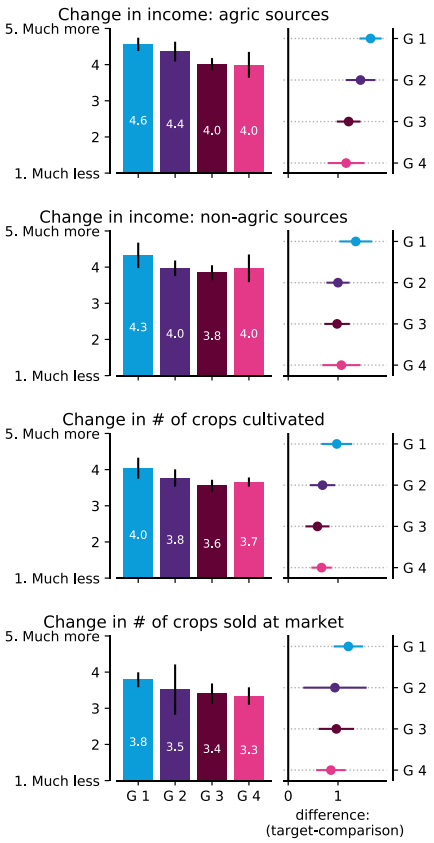
All in all, we thus find that the project is indeed successful in changing peoples motivation, their resilience and farmers' stewardship of the land. The biggest changes occur in people's motivation, followed by stewardship, and finally, resilience. This substantiates the theoretical backbone of the programme which focuses on building farmers intrinsic motivation first, as a basis for further investments in their farm.

Impacts of the PIP-approach on other outcomes

The survey inquired whether farmers observed changes in their income, the number of crops they cultivated, and the number of crops they sold at the markets, over the past three years. Note that this is a subjective evaluation of those changes by the respondents or more accurately their perceived change in income, as opposed to an actual change in income based on the monetary value of a household's total income or consumption.

What we find is that all PIP-participants report that their incomes from both agricultural and non-agricultural sources have increased the past three years. The bulk of PIP-participants reports that their income from agric-sources, the number of different crops that they cultivate, and the number of different crops they sell at markets has increased substantially, over the past three years. All PIP farmers in all generations are more likely to report increases in their incomes compared to non-participants. Earlier generations report (G1 and G2) report larger perceived increases in their incomes compared to later generation (G3 and G4). Although the increases reported by later generations are somewhat smaller compared to the earlier generations, later generations still report that they have a higher income and cultivate and sell more crops compared to farmers who did not participate in the PIP-approach.

Impacts on income and crop yields, by generation



Left: Averages on subconstruct
Right: differences Generation- (matched) comparison (treatment effect)
Thick lines represent 95% confidence intervals

Figure 27 PIP-approach effects on changes (past 3 years) income, crop yields and diversity, by generation.

In addition to the outcomes above[9], we also assess the effect of the PIP-approach on farmers food security situation. The PAPAB-programme, as a whole and the PIP-approach in specific, seeks to increase farmer's food security as its long term impact. The survey included an item where each respondent had to rate whether they have enough food, whether they had barely enough food, or not

[9] Most of these outcomes were initially conceptualized as part of our resilience measurements. However, as some of these are time-series or self-assessments and thereby do not capture a difference at the end of the project we chose to analyse these separately.

enough food at all, for each month, the past year (2019). Figure 28 presents the results[10]. First, we see that overall (throughout the year), the share of people that reports to have insufficient food is higher in the comparison group than among PIP-farmers, throughout the year. Second, the range of farmers who report food shortages throughout the year is substantially broader in the comparison group than among PIP-farmers, indicating greater volatility in food security situation in the comparison group. Third, in some months, the difference between PIP-farmers and non-PIP farmers is not statistically significant (shaded areas overlap in some months). The difference is statistically significant however during the lean season, that is, in the months when all farmers are at their peak of food insecurity or food insecurity is trending upwards (in Oct. Nov. and Dec, the lines do not overlap). Arguably, these are also the months when it matters most. During the harvesting season or the months when general food insecurity is trending downwards, the differences between PIP and non-PIP farmers are not statistically significant.

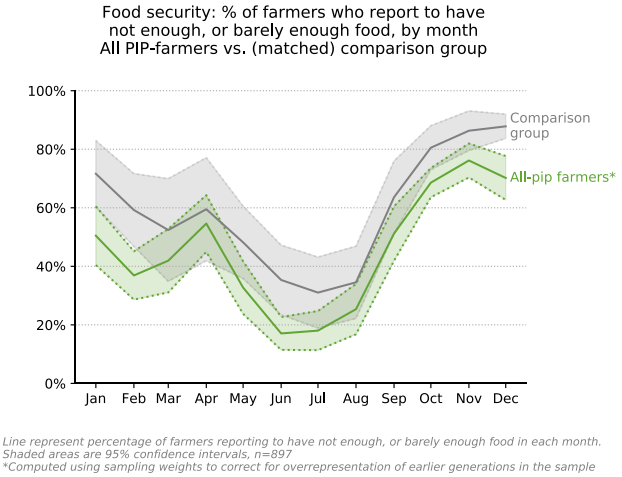


Figure 28 PIP-approach effects on food security throughout the year.

PIP-farmers' food security fluctuates less over the year, and they are especially better able to cope with and mitigate their food insecurity during a context of adversity (the lean season). Thus, these findings are exactly in line with resilience-theory, where increasingly more attention is being focused on people's ability to mitigate and cope with a context of adversity such as the lean season. These results provide strong support for the notion that the PIP-approach is effective in decreasing food insecurity specifically and more generally increasing farmer's resilience.

All in all, this section has demonstrated that the success of the PIP-approach extends beyond increases in farmers motivation, resilience and better stewardship of the land. It indicates that PIP-farmers are reaping the benefits of that increased motivation, resilience, and stewardship in the form of selling more produce, gaining higher incomes and having less shortage of food.

[10] We do not split between generations for reasons of brevity and present the average for all PIP farmers instead. The conclusion presented here are similar in case we do disaggregate between generations.

3.3 RESEARCH QUESTION 2: UNPACKING THE CAUSAL CHAIN

Given the theoretical mechanism presented in section 2.1.2 we expect that the motivation, resilience and stewardship are positively associated with each other. Or in other words, that a more motivated farmer also has higher levels of resilience and is a better steward of the land. In this section, we will dive into research question 2, where we will investigate whether the data gathered fits the theoretical model of increased motivation resulting in more resilience and better stewardship of the land.
It is important to note that research question 2 principally focuses on investigating relationships between the pillars and not on the impact of the PIP-approach on farmers (levels of) motivation, resilience, and stewardship. We have shown in the previous section that PIP-farmers are more motivated, more resilient, and better stewards of the land compared to their non-PIP-peers. However, there is no reason to assume that the relationship between motivation, resilience, and stewardship would be different in a group of farmers that did not participate in the PIP-approach. We also expect that, among non-PIP farmers, those who are more motivated, are more resilient, and better stewards of the land. This means that we will focus our analyses on the whole sample and do not distinguish between PIP and non-PIP farmers.

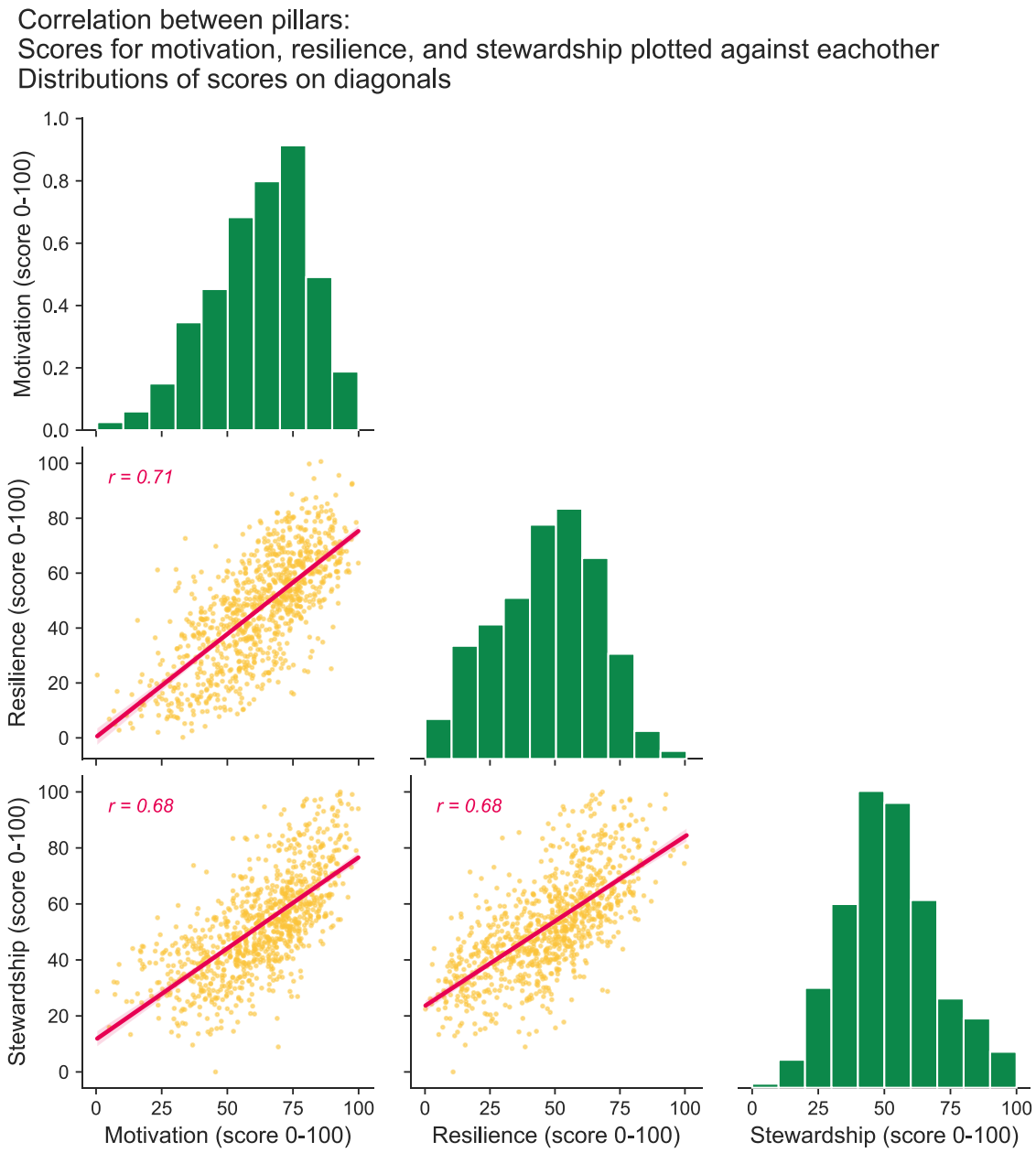
3.3.1 ARE MORE MOTIVATED FAMERS MORE RESILIENT? AND DOES RESILIENCE INCREASE STEWARDSHIP OF THE LAND?

Figure 29 shows all pairwise relationships between the main pillars of the PIP-approach. We use the overall scores for each pillar, as presented in section 3.2 to visualise the relationships between motivation, stewardship, and resilience. The plots on the diagonal show the distribution of scores on a pillar in the whole sample. For instance, the upper left plot (in green) shows the distribution of motivation scores in the whole sample. The plot below plots each farmer's value on motivation (horizontal axis) against that particular farmer's resilience score (on the vertical axis). Each farmer that we have interviewed is represented as a yellow dot. The line in red shows the correlation coefficient, which is a measure that tells us how strongly, in this case, motivation and resilience are associated with each other. The measure ranges between -1 and 1. The value 1 means a perfect positive correlation, the value of -1 means a perfect negative correlation, and the value 0 means that there is no relationship between the characteristics studied at all.

We find that more motivated farmers report higher levels of resilience. There is a particularly strong association between motivation and resilience (correlation coefficient of 0.71). Farmers with high levels of motivation are also more resilient (they cluster in the upper-right corner of the plot). Conversely, those with low motivation, are also shown to have lower scores on our resilience scale. There are virtually no farmers with little motivation, and very high resilience (they would show in the upper left corners of the plot, which is empty).

Moving on the next step in the causal path, the relationship between resilience and stewardship, we find positive and strong relationships as well. Albeit a somewhat weaker association as between motivation and resilience, the strength of the association between resilience and stewardship is 0.68 on a scale ranging from -1 (perfectively negative association) to 1 (perfectly positive association). We thus find that more resilient farmers take more measures to be good stewards of their land.

Additionally, we have also visualised the relationship between motivation and stewardship (the plot in the bottom left corner. We show that higher levels of motivation go hand in hand with more stewardship of the land as well.



r = Pearson's r ranging from -1 perfect negative correlation to 1, perfect positive correlation
all correlation coefficients significant at p<0.01
full sample, n=962

Figure 29 Relationships between motivation, resilience, and stewardship.

When controlling for a myriad of other factors, such as gender, socio-demographics, and region, the relationship presented still hold. These plots only show the simple relationships between the pillars (e.g. solely the relationship between motivation and resilience). Such a relationship might come about as a result of various other alternative explanation. For instance, farmers might be more resilient because they are relatively well-to-do. By controlling for a myriad of other factors, such as gender, socio-demographics, and region, the relationship presented still holds (i.e. the same determinants we have used in our matching algorithms) we rule out such alternative explanations (see Annex 2.2 and 2.3) meaning that the correlations shown here are not spurious. These multivariate analyses, which are a stricter litmus further supports that the relationships between the pillars are as expected and laid out in the theoretical framework behind the PIP-approach.

3.3.2 WHICH ASPECTS OF MOTIVATION ARE ASSOCIATED WITH INCREASED RESILIENCE?

In Figure 30, the correlation coefficients between the sub-constructs for motivation and the pillars resilience and stewardship are presented. Darker green means a stronger positive relationship. Let us look at the first column in that table. This is done to assess whether there are particular aspects of motivation that translate into increased resilience.

The first thing that stands out in figure 30 is that all relationships are positive (and statistically significant). Thus, a farmer's aspirations in life (purpose), an attitude that is geared to learning from others (attitude), the feeling that one is free to make their own choices without pressure from others (autonomy), a supportive household (household support), and fellow villagers (village support), all contribute positively to a household's and farm's resilience.

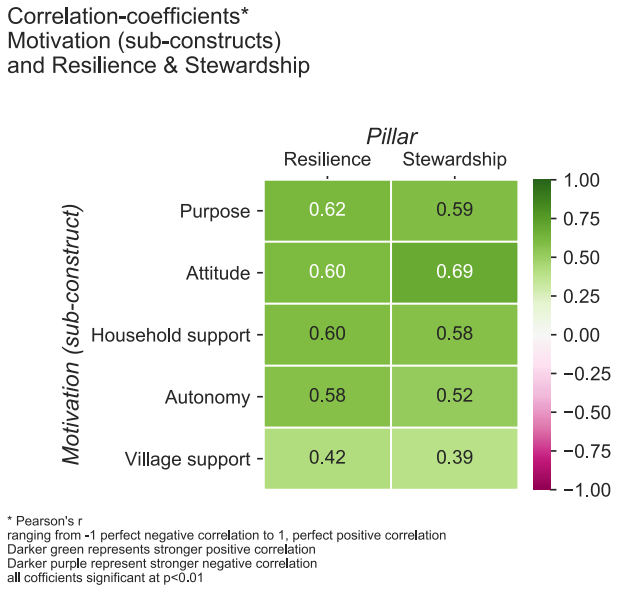


Figure 30 Correlation between sub-constructs for motivation.

Although all associations are positive (and statistically significant), we see that the perceived social acceptance by other villagers, trust and sense of collaboration within the village (village support) seems to be less important than the other indicators in improving resilience. The associations between village support and resilience is the weakest.

The value of the project emphasis on decreasing farmers' dependence on others, and its focus on the whole household as a cornerstone of the approach is shown by the strong correlations between household support, attitude and purpose and resilience. We also see that a supportive household is just as critical (correlation coefficient of 0.60) as farmers autonomy.

We have also plotted the direct relationships between motivation and stewardship for reference. We see a similar pattern. All correlations are positive, and thus as expected given the theoretical foundation of the PIP-approach. Again we find strong correlations for those indicators that capture attributes of the household and collaboration between household members (household support, purpose and attitude) Involving the whole household to devise a vision for the future and implementing these plans does not only contribute to more resilient farming but also to better stewardship of the land.

We do not find any weak, null, or negative relationships for the aspects of motivation studied here. These positive correlations, for all sub-constructs on resilience and stewardship, support our finding that the PIP-programme is working to increase farmers motivation by focusing on the right aspects of motivation. This also means that none of these aspects can be ignored because they all induce resilience and stewardship. This endorses the notion behind the PIP-approach that it is indeed necessary to take a holistic approach, working with the whole household together on multiple aspects of motivation simultaneously and comprehensively.

3.3.3 WHICH ASPECTS OF RESILIENCE ARE ASSOCIATED WITH INCREASED STEWARDSHIP?

In Figure 31, the correlations between the aspects of resilience and overall stewardship are presented in the same manner as in figure 31. We see that all aspects of resilience contribute positively to the stewardship of the land. The associations are somewhat weaker compared to those between motivation and resilience. We again find support for the path as posited in the PIP-approach's foundational theory. The same lines of reasoning offered in the previous section hold here as well. All relationships turn out to be positive and significant. We do not see negative relationships or relationships that are absent. Therefore we cannot exclude any of these aspects of resilience in the project's training; all aspects play a role in increasing farmers stewardship of the land.

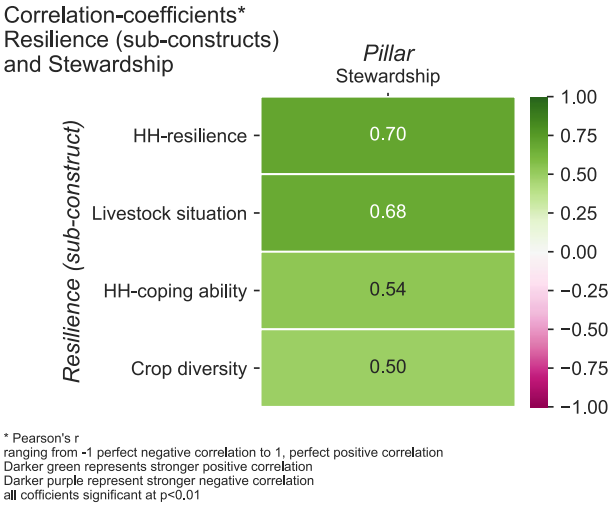


Figure 31 Correlation between sub-constructs of resilience and stewardship.

Household resilience and livestock situation play the most important role in increasing stewardship, crop diversity is the least important. An interesting note here is that we have seen that the livestock situation (which constitutes the different types of livestock respondents hold and whether they have enough resources (fodder) available for their livestock) is an important determinant of overall resilience (see Table 12 in the Appendix). Recall that we do find an impact on the livestock situation for generation 1 and generation 2, but not for generation 3 and 4 and that generation 4 but particularly generation 3 is less likely to hold livestock compared to the earlier generations. The strong correlation between livestock situation and stewardship stipulates the importance of investing in livestock and moving from a crop-production only to a mixed farm. Those farmers with a wider diversity of livestock and a better ability to feed for their livestock are also better stewards of the land. Livestock diversity and having the resources to properly care for one's livestock is thus not only an important aspect of resilience, but this effect also cascades through towards better stewardship.

3.3.4 SYNTHESIS

Is participation in the PIP-approach associated with increased motivation, resilience and stewardship?

The previous sections have shown that motivation, resilience, and stewardship are all positively related. We have not made an explicit connection between participating in the PIP-project and the relationships between the pillars. Given the marked differences between PIP and non-PIP farmers shown in section 3.2, one would expect that making a PIP plan and following up on these plans would increase farmers motivation, their resilience, and their stewardship. The hypothesis we test explicitly now is that those farmers who have been participating longer in the PIP approach and followed up on more of their plans are more motivated, more resilient and better stewards of the land. We do this by investigating whether the extent to which PIP-plans have come to fruition (the PIP-plan completion rate) is associated with motivation, resilience and stewardship.

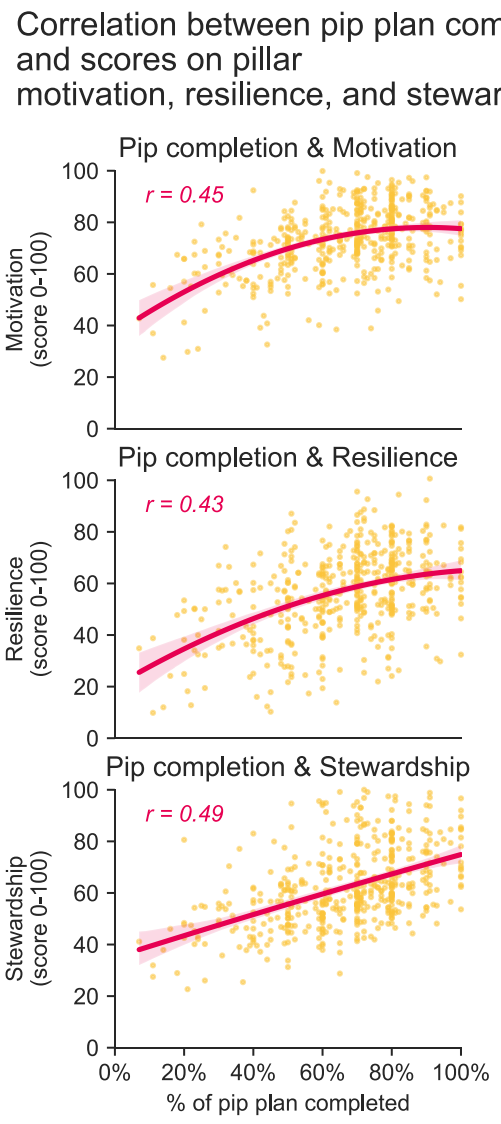


Figure 32 Correlations between PIP-plan completion, motivation, resilience, and stewardship

Figure 32 shows the relationship between the completion of the PIP-plan (horizontal axis) and the farmers score on each respective pillar.

We find that farmers that have progressed further in implementing their PIP-plans are more motivated (top plot). Similarly, we find that further completions of the PIP-plan breeds resilience (middle plot.) Likewise, those who have implemented a larger share of their plans report higher levels of stewardship (bottom graph).

Note that the relationships between PIP-completion rate and all pillars are distinctly curvilinear[11]. The increase in motivation, resilience and stewardship is greater at lower rates of PIP completion (<50% of the plans implemented) compared to higher rates of PIP completion. This shows that motivation, resilience and stewardship increase the fastest directly after entering the PIP-project. Moreover, there are so-called ceiling effects. Once, say 50-60% of the PIP plan is implemented, there are relatively fewer gains in people's resilience and motivation of further implementing the PIP-plan.

All in all, we see that implementing the plans as developed in the PIP-approach are associated with higher levels of motivation, resilience, and stewardship. The benefits of the PIP-plan for particularly increasing resilience, motivation,

[11] We have estimated linear relationships between these characteristics as well but found that estimating polynomials (in this case quadratic terms) describe these relationships better. This was tested by comparing the model fit for regressions using a linear term provide against models that included a quadratic term. In all cases this provided a statistically significant better model fit in terms of its explained variance on the relevant pillar. This means that the increases in pillars scores do not progress linearly across each 'step' in PIP-completion rate. The (marginal) increases, particularly in motivation and resilience, are much bigger at the beginning of the PIP-implementation compared to later stage.

grow quicker in the earlier stages and seem to wear off in later stages where motivation and resilience reached a ceiling. Despite these ceiling effects, these relationships further corroborate our finding that the PIP-approach is effective.

4 CONCLUSION

In this study, we have assessed two main questions. The first focused on the impact of the PIP approach. We assessed whether PIP-farmers displayed more motivation, are more resilient, and better stewards of their land compared to non-PIP farmers. The second questions centred on causal paths presented in section 1.1. the theoretical backbone behind the PIP-approach. We assessed whether the PIP-approach follows the paths as theoretically expected by diving deeper into the relationships between the main pillars of the PIP approach.

We started by closely looking at who participates in the PIP-approach by comparing all generations of farmers on a set of background characteristics. Looking at the composition of PIP-farmers, we find that the first generation of farmers is relatively well to do. They have more land and livestock at their disposal, are higher educated, and more often have mixed farms. They are relatively well to do both in comparison with non-PIP-farmers as well as in comparison with later generations. Results show that the socio-demographic profile of later generations is more mixed. Particularly the third generation seem to be relatively poorer compared to the earlier generations. This is a result of the targeting strategy of the PIP-project and sequence in the competitions and training where the first generation are selected based on a more innovative progress-driven mindset.

As mentioned, such a targeting strategy can be a double-edged sword. On the one hand, having the community choose the farmers that participate as the first generation likely takes in those farmers where other farmers look up to . These first generation farmers receive acclaim from other farmers in the village, and can then guide and lead other farmers by example. On the other hand, this is also increases the selectiveness of the targeting mechanism in the first generations which risks making the project less inclusive by creating an entry barrier for poorer farmers.

We see that the PIP-approach, with its competitions, does broaden and includes farmers that are lower on the socio-economic ladder in each further generation. The third generation is particularly interesting in this respect, as a large share of this group consists of third generation farmers in so-called initial collines. Initial collines are the collines where the PIP approach started with the ‘Paysan Innovateurs’. Farmers of generation 1, generation, 2 and generation 3 reside in these initial collines (whereas in extension collines we only see generation 3 and generation 4 farmers). The mechanism where the first generation of relatively well to do farmers stimulate poorer farmers to join the project through leading by example, or vice-versa, a selective non-inclusive sequential targeting, would be most prominent in these initial collines. We find that in the initial collines, the third generation is often the last group of farmers in a colline to enter the programme through its demonstrations and competitions, and also the poorest group of farmers. Although there might be an entry barrier for these poorer farmers in the first stages of the project (i.e. being included in generation 1 or generation 2), relatively poor farmers are able to enter the programme in later stages. This finding corroborates our conclusion that the targeting strategy is effective and inclusive. The sequential targeting works. The more motivated and relatively well off farmers receive the sought after acclaim through leading by example without leaving out poorer farmers in later generations.

The impact on motivation seems to be bigger for earlier generations compared to later generations. This pattern holds throughout virtually all the indicators studied. The first generation has the highest levels of motivation, stewardship and resilience, and they also differ to a larger extent with their comparison group (which is composed of a group of similarly well-to-do farmers). Generation one’s motivation score is about 1.5 times higher compared to a similar group of farmers who do not participate in the PIP-approach. The impact on motivation is apparent throughout all generations. Even the third and fourth generation of farmers which have been participating in the programme for a shorter period

of time and where especially generation 3 comes from a comparatively lower point of departure report levels of motivation that are more than 1.3 times higher as those who do not participate. Generally, the project is very effective in increasing farmers motivation.

PIP-farmers report higher levels of household resilience, and households feel more able to cope with exogenous shocks compared to non-PIP farmers. The project is thus effective in changing these attitudinal aspects of resilience. However, the PIP-approach is less effective in changing the physical aspects of resilience, particularly in increasing crop and livestock diversity. We do see that Generation 1 and Generation 2 grow more diverse crops and have a wider diversity of livestock but that this is not the case for Generation 3 and Generation 4 (benchmarked against their comparison groups). Adaptations in and around the farm itself (growing a wider variety of crops) likely need a longer time to achieve. Moreover, diversification of the whole farm itself, that is departing from crop-only farms to mixed farms that grow a wide variety of crops and hold different animals, so that farmers can rely on a wider variety of products to sell is of paramount importance to farmers' resilience. Moreover, we demonstrate that having a diverse set of livestock and the ability to care for these animals is an important determinant for good stewardship of the land as well. Increasing livestock diversity and shifting towards a mixed farm might be more difficult for poorer generations of farmers (Generation 3) who are less likely to possess (financial) resources necessary to invest in productive assets such as livestock that increase their resilience.

Although the PIP-approach is not effective in improving all sub-aspects of resilience when taking a more comprehensive look at the pillar as a whole, the results show that the PIP-approach affects farmers' resilience positively, for all generations of farmers.

Regarding stewardship we find that PIP-farmers do have more knowledge about a variety of farming practices, know better how to use fertilizer and are more aware of changes in soil, vegetation and more generally their land compared to non-PIP farmers. They thus have more knowledge about why and how to implement various farming practices.

More knowledge does not always translate into implementing a different set of farming practices. We find that PIP-farmers are more likely to combine chemical and natural fertilizers on their plots, rotate crops, dig trenches and are more likely to plough on the contourlines compared to non-PIP farmers. Most of these practices (except crop-rotation schemes and their implementation) are already relatively common practice, among non-PIP farmers as well. We do not find an impact of the PIP-approach on other farming practices. PIP-farmers are not more likely to build gully control structures or stonebunds on their land. Likewise, the use of mulches, cover crops, and staggered row planting is not more common among PIP-farmers across all generations compared to non-PIP. We do acknowledge that some of these practices are heavily dependent on the idiosyncrasies of each plot of land (e.g. gully control), and might thereby be rare by definition. Although PIP-farmers are more conscious about the why and how of these farming practices, they do not differ in the implementation of most farming practices studied here. The array of farming practices that PIP-farmers implement is not markedly distinct from the set of farming practices that non-PIP farmers use. Rather, PIP-farmers are slightly more likely to implement those practices that are already relatively common among non-PIP farmers as well.

Nevertheless, when taking an all-encompassing look at the stewardship pillar as a whole, we still find strong impacts of the approach; PIP-farmers are thus better stewards of their land compared to non-PIP farmers.

We find strong impacts on virtually all opinions, self-assessments, and attitudes across all pillars. We do find impacts on some but not all behavioural aspects of the resilience and stewardship. This should be interpreted with reference to the fact that development programmes are often plagued by attitudes not directly and unambiguously translating into a different set of behaviours. This disconnect also plays a role in some aspects of the PIP-approach, particularly when it comes to more structural aspects of resilience (e.g. investing in (more diverse) livestock) and farming practices that foster stewardship. Nevertheless, our analyses at pillar level, that is taking all attitudinal and behavioural

aspects into account demonstrates that PIP-farmers are indeed more motivated, more resilient, and better stewards of their land.

The findings also show that the success of the PIP-approach extends beyond these increases in farmers motivation and resilience and stewardship of the land. The analyses on the set of outcomes that measured perceived changes in sales and income show that participants reap the benefits of that increased motivation, resilience, and stewardship in the form of selling more produce and are ultimately gaining higher incomes. Likewise, these effects are stronger for earlier generations of farmers compared to later generations of farmers. Moreover, PIP-farmers are more food secure in general and especially at times when food security matters most (i.e. lean(er) seasons).

Overall our data confirms the core claims and expectations that flow from the theoretical grounding of the PIP approach. Indeed, motivation is associated with more resilience and more stewardship. This is confirmed through the analyses at pillar-level. The more detailed deep-dive into sub-constructs revealed that also all sub-constructs are strongly and positively correlated with the next pillar on the causal chain. This again highlights the value of the holistic and comprehensive nature of the approach that involves the whole household and combines a wide variety of aspects regarding motivation, stewardship and resilience.

When assessing the relationships between the pillars and the sub-constructs, we find strong and positive relationships between the pillars as theoretically expected. We provide more rigorous evidence that the PIP-approach is responsible for increasing motivation stewardship and resilience because we see that farmers who have progressed further in the implementation of their plans, display higher levels of motivation, stewardship, and resilience. Moreover, we find that the most gains, especially in motivation and resilience, are made in the early stages of the implementation of the PIP-plans. The marginal increases in motivation and resilience are smaller in the later stages of implementation.

We have presented the findings in the report following a rather strict causal chain where motivation fosters resilience, which in turn leads to stewardship. We do acknowledge that these relationships are not that simple and clear-cut. They are all part of a well-tossed bowl of spaghetti of correlations that is very hard to disentangle. The reasons for strictly adhering to this is methodological and to provide a lens through which to view the results. In some cases, the causality may be reversed. An increase in stewardship might also motivate people. Equally, better stewardship might lead to more resilience. The design chosen does not allow for the strictest tests on whether resilience or stewardship comes first. Nevertheless, the results confirm the relationships or associations as theoretically set out, in both rudimentary as well as more rigorous empirical tests.

All in all, the data provide strong evidence that the PIP approach generates proud farmers who realize that their land is their main asset, and who feel able and intrinsically motivated to invest in their farms, to become resilient farmers and good stewards of their land.

5 RECOMMENDATIONS

The results demonstrate that the support from the whole household plays an important role in the PIP-farmers transformation to building more resilient livelihoods and become better stewards of the land. Thus, do involve, and keep involving the whole household in drawing out the PIP-plans and in seeking ways to build a more resilient and sustainable future for farmers.

This study also shows that it is necessary to take a more encompassing and holistic approach that focuses on farmers intrinsic motivation as opposed to highly specialized and ad-hoc pieces of training that focus on extrinsic sources of motivation such as financial rewards. The theoretical backbone of the PIP-approach is grounded in this multifaceted approach, and its validity is supported through our analyses of relationships between the pillars and their sub-constructs. The interplay between and within pillars is as expected. This leads us to recommend to continue this multifaceted approach as it is currently being implemented, instead of choosing for a more narrowly defined approach, by cherry-picking or choosing to single out a narrow sub-aspect to work on in the PIP-activities.

We do not find evidence that PIP-farmers diversify their crops and livestock to a larger extent than non-PIP farmers. PIP-farmers are more resilient farmers because they are better able to cope with, and mitigate a variety of sudden adverse events (shocks) that negatively affect their livelihoods, but the training and activities do not lead to a diversification of crops and livestock (products). We, therefore, recommend paying more attention to tools and skills that can help PIP-farmers to diversify their production of both crops and livelihood products. Livestock diversification is especially important here because this is not solely important for increasing resilience, but this is also an important determinant of stewardship.

Regarding stewardship, we find that PIP-farmers are generally better stewards of their land but that they do not implement a markedly distinct set of farming practices compared to non-PIP farmers. The recommendation here is to more thoroughly investigate the barriers PIP-farmers face in the implementation of these not-so-common practices that foster stewardship.

We have found the targeting strategy to be very effective as well. The data provides support for the notion that initially targeting relatively well-to-do farmers makes them lead by example and engages other, relatively poorer farmers in later generations. The competitions and demonstration do reach poorer farmers and does not lead to the project to become exclusive and only for the well to do.

We also see that the most gains in terms of motivation, resilience, and stewardship are made early on in the programme (and in the early stages of implementation of the PIP plan) and that the marginal gains are smaller in later stages. The effects stall at the ceiling of motivation and resilience. Therefore, we recommend having a quicker follow up of PIP-generations compared to the current situation (after a period of 2- 3 years) wherever possible as most gains are made in these early stages.

Finally, the PIP approach is effective on the vast majority of indicators studied here, most notably on farmers motivation. The analyses confirmed that relationships set out in the theoretical backbone of the PIP-approach hold empirically after rigorous testing. PIP farmers are generally more motivated, more resilient and better stewards of the land compared to non-PIP farmers and the effects of the project extend beyond this. PIP farmers report that their income and sales have increased the past years, and more so compared to non-PIP farmers. Moreover, PIP-farmers are more food secure in general and more food secure in times of duress.

Overall, this study shows that the PIP-approach is very effective in increasing motivation, resilience, stewardship, and even beyond, its theoretical backbone and assumptions are well-grounded and hold empirically after rigorous testing. Therefore our final recommendation is plain and simple. Continue to implement the PIP-approach.

6 REFERENCES

Hair JF, Tatham RL, Anderson RE and Black W (1998) Multivariate data analysis. (Fifth Ed.) Prentice-Hall:London.

Holling, C. S. (1973). Resilience and stability of ecological systems. *Annual review of ecology and systematics*, *4*(1), 1-23.

Kessler, C.A. & van Reemst, L.I. 2018. *PIP Impact report*; Wageningen, Wageningen Environmental Research.

Lance, C.E., Butts, M.M., & Michels, L.C. (2006) The Sources of Four Commonly Reported Cutoff Criteria: What Did They Really Say? *Organizationsal Research Methods,* 9:2 (202-220). https://journals.sagepub.com/doi/abs/10.1177/1094428105284919

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American psychologist, 55*(1), 68.

ANNEXES

ANNEX 1: SAMPLE SIZE OVERVIEW

Table 1: Detailed overview of number of respondents per colline and generation

Province	Commune	Colline	PIP G1	PIP G2	PIP G3	PIP G4	Comparison
Bujumbura	Kanyosha	Musugi				14	
		Ruvumu	13				
	Mubimbi	Gisagara		15			
		Kiziba			12		
	Mugongo-manga	Gisarwe					60
		Mwura					60
	Nyabiraba	Bubaji				15	
		Kizunga				13	
		Nyabiraba		16			
Cibitoke	Buganda	Kansega					60
	Bukinanyana	Kibaya					60
	Mabayi	Buhoro	13		12		
		Gitukura			12		
	Rugombo	Kagazi	12		23		
		Rugeregere				14	
Makamba	Mabanda	Kibago					59
	Makamba	Canda		16	13		
		Nyankara				14	
	Nyanza-Lac	Biniganyi	17	12			
		Mugumure		16			
		Mukubano				15	
Muyinga	Buhinyuza	Butihinda					60
	Muyinga	Gatongati 1				14	
		Murama	13		12		
	Mwakiro	Gahekenya		15			
		Gahemba	13	15	12		
		Rukanya				14	
Rumonge	Burambi	Buhinyuza				14	
		Gatobo		15			
		Rwaniro	14				
	Buyengero	Karambi					60
	Muhuta	Gitaza					54
	Rumonge	Gatete			12		
		Muhanda				14	
		Murambi	13		12		
Total			**108**	**120**	**120**	**141**	**473**
Overall Total							***962***

ANNEX 2: METHODOLOGY

2.1 SCALABILITY: SURVEY-ITEMS, SUB-CONSTRUCTS, AND PILLARS

We acknowledge the multifaceted nature of motivation, resilience, and stewardship (the pillars) by conducting analyses on the separate facets (or sub-constructs) within each pillar. Here, we can see what the effect of the PIP-approach is on, for instance, household support separately. Subsequently, we assessed the effectiveness of the PIP-approach on the pillar as a whole. This section describes how the items in the questionnaire are related to sub-constructs, and how sub-constructs relate to the pillars. It furthermore presents the scaling procedures used to come to theoretically and empirically meaningful measures of motivation, stewardship, and resilience.

Scaling techniques

Rationale

The objective of using various scaling techniques that merge items into composite indicators at sub-construct (e.g. household support) and pillar level (e.g. motivation) is twofold.

First, our objective is to investigate whether the questions in the questionnaire are related to each other and measure the same overarching sub-construct. During the questionnaire design phase, the pillars and sub-constructs were operationalised into actual questionnaire items. The techniques employed here offer an empirical justification of the initially envisioned operationalization of sub-constructs into questionnaire items.

By scrutinizing the association between various questions, we investigate whether the model initially envisioned during the questionnaire design stage fits the data collected. Thus, for instance, for the sub-construct purpose (in the pillar motivation). The scaling techniques tests whether the responses to a certain item (e.g. “Do you and your household generally value your life and the things that you do?”) are associated with the responses to other questionnaire items that we expect to be indicators of the sub-construct purpose. By extension, they thus provide a test whether certain questions (or even sub-constructs as a whole) are relevant to incorporate the measurement of a certain sub-construct. After all, in case responses to an item such as “Do you and your household generally value your life and the things that you do?” are not related to those other items that measure the sub-construct purpose, this means the non-related item is not a good indicator of the overarching concept.

Often, the concepts necessary to study cannot be directly observed (e.g. motivation is not something that can directly be counted, or ‘seen’) and are thus latent in nature. The theoretical backbone of measuring such attitudes is rooted in item response theory. If we cannot directly observe or ‘count’ a concept such as motivation, we ask a series of related questions or statements in a questionnaire. By doing so, the survey transfers this concept into multiple manifest measures (manifest in the sense that we can directly observe and quantify a respondents’ answer to a single statement). Next, we can determine to what extent the responses on these items are associated with each other. When responses on the items are associated with each other (high scores on one statement correspond with high scores on another statement, and vice-versa) there is strong evidence to suggest that these manifest measures are indeed measuring the same underlying latent concept.

Second, scaling procedures provide the benefit of reducing the dimensionality of the data. If one used each item separately in an analysis, this would result in an accumulation of measurement error associated with each individual item. Consequently, it would be impossible to separate the signal from noise both in the statistical estimation techniques we use and, more importantly, in the subsequent interpretation of these effects. The scaling techniques that we employ allow us to generate a single score for each sub-construct for each respondent in the survey based on their responses to whole batteries of items.

For concepts that are solely attitudinal in nature (e.g. motivation and all sub-constructs), we employed factor analysis a means to come to appropriate scales for each sub-constructs and pillar. Factor analysis requires the underlying items to be strongly correlated with each other and is appropriate to test whether the items (questions in the questionnaire) actually refer to the same underlying latent concept.

The pillars resilience and stewardships contain items that are not attitudinal or latent in nature but rather phenomena which can directly be observed. For instance, when measuring crop diversity, the number of annual and perennial crops is unambiguously 'countable'. The pillars resilience and stewardship consist of a combination of such counts and latent or attitudinal measurements (cf. the number of different crops cultivated, and the statements about attitudes used to measure household coping ability). For the latent measurements in these pillars we do implement factor analyses to test whether the items used are a representation of a single concept (at sub-construct level) (i.e. do the items actually measure someone's knowledge and awareness about the use of the commons).

To construct the overall resilience and stewardship scores, we are not determining whether each questionnaire item is a representation of some underlying latent construct because some practices of interest are directly observable (e.g. the interviewer can see for instance gully control structures). Instead, for overall stewardship and resilience, we are merely interested in generating an easy to interpret single overall score boils the long list of items down to a single metric. For resilience and stewardship, we use component analysis to come to these composite indices (the overall stewardship and resilience scores) and generate the weights for each component or sub-construct. Although the guiding principles between factor and component analyses are similar, the difference is that factor analyses explicitly test whether items represent a certain latent construct by estimating which share of the variance is communal among all items and which share is unique to a single item. Component analyses do not rely on this decomposition of variance into unique and communal but instead assumes that each item fully contributes to the index that needs to be constructed.

In these factor and component analyses we follow Hair et al., (1998) to determine cut-off values for in or exclusion of an item in a certain scale (a factor loading cut-off of 0.3 suffices but values higher than 0.7 indicate strong scaleability for an item). For those sub-constructs that are attitudinal, we additionally present Cronbach's α. This is a measure of internal consistency ranging between 0 and 1 and can be colloquially interpreted as the average covariance between the items and the total scale-score constructed. Following Lance et al. (2006) Cronbach's alpha higher than 0.7 indicate high internal consistency and reliability, whereas values higher than 0.5 indicate moderate reliability (see: Hinton et al., 2004).

Scale construction

Scores on a sub-construct

To generate values for sub-constructs, we use the scaling techniques on the relationship between item and sub-construct. The questions we seek to answer is for instance: Are the responses on the item "Do you and your household generally value your life and the things that you do?" sufficiently associated with the other items used in the sub-construct purpose?". If so, we construct a composite indicator for purpose that includes the item above. In case the empirical tests fail, given cutoffs describe earlier, we exclude the item from the purpose-scale. In case the combination of questionnaire items proves to be a reliable indicator of a sub-construct, each respondent is assigned the average value of the relevant items as a score on a certain sub-construct[12].

[12] In cases where answer categories are not similar within a scale, for instance one question has answer categories ranging from 1-3 whereas others used standard likert-scale answer categories from 1-5 we have rescaled all items to have a consistent range to avoid a single item to disproportionally affect the composite scale solely as a result of having more (or less) answer categories.

From sub-construct to pillar

Next, we move one level higher and use the scaling techniques to assess the relationship between sub-construct and pillar. The question we seek to answer remain s the same. For example: Are the response on the relevant scale for the sub-construct purpose sufficiently associated with the responses on the other sub-constructs(e.g. autonomy, attitude, etc.) in the relevant pillar motivation.

For the step going from sub-constructs to pillar, we take a slightly different approach in calculating the scale-scores. We construct a weighted sum of all sub-constructs within a pillar, where factor loadings of the relevant sub-constructs determine these weights. Next, we normalize these weighted sums to values between 0 and 100 for ease of visualisation and interpretation. In practice, this means that each farmer has a motivation, stewardship, and resilience score between 0 - lowest and 100 - highest.

The next sections describe these scaling procedures, first, item to sub-construct, and then from sub-construct to pillar.

Pillar 1: Motivation

Items to sub-constructs

Purpose

Table 2: Sub-construct Purpose

Sub-construct	*Questionnaire item*	*Answer categories*	*Factor loading on sub-construct purpose*	*Composite measure used*
Purpose	Do you and your household generally value your life and the things that you do?	Not at all, never (1) – Of course, no doubts (5)	0.7143	Average on 5 item scale ranging from 1-low purpose to -5- highest purpose.
	Do you feel proud about the life you and your household have here in the community?	Idem	0.7154	
	What do you think will be the condition of your household in 5 years' time?	Much worse (5) – Much better (5)	0.7053	
	Describe the plans or aspirations your household has for the near future to improve the conditions of the household	No plans and aspirations (1) – Lots of plans and aspirations (5)	0.6215	
	Describe concrete actions that your household has taken the past year to improve the conditions of the household	No concrete actions done – Lots of concrete actions done (5)	0.6530	
Cronbach's α= 0.8127				

The sub-construct of purpose is a combination of five questionnaire items, shown in table 2. Each item has a factor loading above 0.6, which is far above the threshold of 0.3 (see above). Cronbach's alpha of 0.8070 indicates the purpose scale to be reliable.

Purpose initially included one extra item, which was "*Are you (and your household) willing to stay and live here over the coming 10 years?*". First, on a more conceptual note, this item is more closely related to peoples intentions to stay where they currently are. Or, if you will, this would be akin to the inverse of a migration intention. Empirically, the factor loading (-0.0150) is below the threshold, showing that item is not consistent with the others, and thus should be excluded from the scale.

Autonomy

The six questionnaire items in table 3 below measure the sub-construct of autonomy. All items have factor loadings above 0.5, indicating high internal consistency. The high Cronbach's alpha of 0.8106 indicates the constructed scale for autonomy to be reliable. All items formulated during the questionnaire design stage that were expected to be indicators of autonomy are retained in the final scale.

Table 3: Sub-construct Autonomy

Sub-construct	*Questionnaire item*	*Answer categories*	*Factor loading on sub-construct Autonomy*	*Composite measure used*
Autonomy	Do you and your household feel free to make your own choices about the households future; i.e. without pressure or force from others?	Not at all, never (1) – Of course, no doubts (5)	0.5156	Average on 6 item scale ranging from 1-low autonomy to -5- highest autonomy.
	Do you and your household feel that you generally act and do things in daily life according to your own interest and desires?	Idem	0.6086	
	Do you and your household feel able to improve the quality of your life without depending on others?	Idem	0.7637	
	Do you and your household find it easy to start new activities by yourselves?	Idem	0.7238	
	Do you and your household feel that you are generally in charge of the situation in which you live?	Idem	0.5767	
	Can you and your household generally manage all the responsibilities that you have in life?	Idem	0.6474	
Cronbach's α= 0.8106				

Attitude

Table 4: Sub-construct Attitude

Sub-construct	*Questionnaire item*	*Answer categories*	*Factor loading*	*Composite measure used*
Attitude	Are you and your household always busy to learn new things on how to improve the farm?	Not at all, never (1) – Of course, no doubts (5)	0.5572	Average on 6 item scale ranging from 1-low attitude to -5- highest attitude.
	When you see changes on other farms, would you then ask the owner what s/he has done to learn from this?	Not at all, never (1) – Always, whenever possible (5)	0.7498	
	How often do you share your knowledge and experiences with others?	Idem	0.7644	
	When you have a problem on your farm, do you often ask others for their opinion or advice?	Idem	0.7764	
	Please describe which new practices/tools (any innovation!) you have tested or started	Total nr of new practices tested/used on the farm, rescaled (1-5)	0.6837	

	to use on your farm / in the household the past year?[13]			
	If you improve something on your farm or in your household, does that make you feel proud?	Not at all, never (1) – Of course, no doubts (5)	0.3305	
Cronbach's α= 0. 7995				

The six questionnaire items in table 4 together measure the sub-construct of attitude. All factor loadings are above the appropriate cut-off value of 0.3. The high Cronbach's alpha of 0. 7995 indicates the constructed scale for attitude to be reliable. All items are retained in the final composite scale, which is the average value on the responses for the relevant items.

Table 5: Sub-construct Household Support

Sub-construct	*Questionnaire item*	*Answer categories*	*Factor loading*	*Composite measure used*
Household support	How would you describe the extent to which members of your household collaborate on the same objectives?	Bad (1) – Very good (5)	0.5238	Average on 5 item scale ranging from 1-low hh-support to -5- highest hh-support.
	Who is doing usually the planning of agricultural activities within the household?	Wife / husband only (1) – Whole household together (5)	0.3969	
	In your household, do you generally understand each other well when it comes to planning of issues on the farm and within the household	Not at all, never (1) – Of course, no doubts (5)	0.5970	
	Does your household usually have enough access to labour to do what you want?	Idem	0.5250	
	Does your household usually have enough money to do what you want?	Idem	0.5581	
Cronbach's α= 0.6564				

The sub-construct of household support is a combination of the five questionnaire items shown in the table above. As can be seen, each item has a factor loading above 0.3, which indicates internal consistency across items. A single item was removed. This was *"How often are there conflicts between the members of your household that are a concern to you?*". As the factor loading (0.1402) was below the threshold, this item is not consistent with the others, and thus should be excluded from the scale. Cronbach's alpha (0.6564) indicates the constructed scale to be moderately reliable.

Table 6: Sub-construct Village Support

Sub-construct	*Questionnaire item*	*Answer categories*	*Factor loading*	*Composite measure used*
Village support	Do you generally feel valued/respected by the members of your community?	Not at all, never (1) – Of course, no doubts (5)	0.4929	Average on 5 item scale ranging from 1-low village-support to -5- highest village-support.
	Do you generally trust the people in this village?	Idem	0.4882	
	Imagine that someone here asks you to lend some money, do you usually give it, knowing that it will be returned to you any time?	Idem	0.5368	

13 This item is treated as latent for ease of analyses and because this involved a judgment call from the enumerator on the practices explained by the respondent.

	In case of conflict, are these in general pacifically solved between the villagers?	Not at all, never (1) – Often (5)	0.3871	
	Do you agree that in this village people generally have the same vision?	Strongly disagree (1) – Strongly agree (5)	0.3825	
Cronbach's α= 0.6564				

The sub-construct of village support is a combination of the five questionnaire items shown in the table above. Each item has a factor loading above 0.3, which indicates internal consistency across items. Initially, the scale included two extra items, which were "*Are there conflicts between the villagers in this village that are a concern to you?*" and "*Does your household have any conflicts with other villagers that are a concern to you?*". As the factor loadings (respectively 0.1045 and 0.2522) were below the threshold, these items are not consistent with the others, and thus should be excluded from the scale. Cronbach's alpha (0.5855) indicates the constructed scale to be moderately reliable.

Motivation: Sub-construct to pillar

The previous section demonstrated the scaling from item to sub-construct and assessed whether items can be merged together in a single sub-construct. This section moves to the level of the pillar. We apply the same procedure to determine whether sub-constructs can be scaled in to a single pillar to ultimately arrive at a motivation score ranging from 0-100.

Table 7: Pillar 1 Motivation

Concept	*Sub-construct*		*Questionnaire items*	*Factor loading*	*Measure used in overall motivation score*
Motivation	Purpose	Measures the extent to which farmers feel they have a purpose in life, their aspirations and have made concrete plans and actions in support of realising these aspirations	5 item scale, average between 1-5 see Table 2 for items	0.7804	Weighted sum of values on sub-construct (weights determined by factor loading) normalized to values between 0-low motivation, and 100 – high motivation.
	Autonomy	Measures the extent to which households feel they can make their own choices in life, are in charge of their own life with and can manage their responsibilities without depending on others.	6 item scale, average between 1-5 see Table 3 for items	0.8129	
	Attitude	Measures the extent to which farmers and their households have an attitude geared towards learning from others, whether they actively seek and implement new farming practices. and have an actively learning attitude	6 item scale, average between 1-5 see Table 4 for items	0.7541	
	Household support	Measures the extent to which household collaborate to achieve shared objectives, and whether they have sufficient access to labour and money to achieve these objectives.	5 item scale, average between 1-5 see Table 5 for items	0.7163	
	Village support	Measures the households perceived support and standing in the village and the levels of cooperation and shared vision within the village as perceived by household.	5 item scale, average between 1-5 see Table 6 for items	0.5833	
Cronbach's α= 0.8578					

The pillar of motivation combines the five sub-constructs purpose, autonomy, attitude, household support, and village support as described in the previous section. Table 7 demonstrate that the factor loadings for each sub-constructs are far above the threshold, indicating internal consistency across the sub-constructs. Cronbach's alpha (0.8562) indicates the constructed scale for motivation to be reliable. The score on the pillar motivation is constructed by a weighted sum (value on subconstruct multiplied by the factor loading). For ease of interpretation and visualisation, this weighted sum is then normalized so that the highest motivation score is 100, and the lowest motivation score is 0.

Pillar 2: Resilience

Items to sub-constructs

The pillar resilience consists of 4 sub-constructs: crop diversity, livestock diversity & situation, household resilience, and household coping ability. The pillar resilience consists of a combination of latent and directly observed measures. For the pillar resilience, we employed component analysis as opposed to factor analysis as outlined in previous sections.

The resilience measurement used here differs somewhat from measures for resilience used in previous research to study the effectiveness of the PIP-approach as well as from the initially envisioned conceptual framework (cf. Kessler & van Reemst, 2018). During the initial operationalisation phase, we expected income diversity to be an essential indicator of resilience. A household that relies on a more diversified range of income-sources would indeed cope better with adverse events because losses incurred from for instance a lower yield during the harvesting season could be offset by incomes from a different, non-agric source. The number of households in the sample that have income sources other than those directly related to agriculture is extremely small. Almost all (98%) respondents are farmer, and the median share of income derived from agric-related sources is 100%. Income diversity does not discriminate, in this particular sample because virtually everyone relies on agric-income sources. We thus exclude this sub-construct from our analysis and instead focus on these sub-constructs (crop-diversity, livestock situation, household resilience, and coping ability) which do display variance across the sample.

Crop diversity

Table 8: Sub-construct Crop Diversity

Sub-construct	*Questionnaire item*	*Answer categories*	*Measures used*
Crop diversity (cultivated & crops sold)	Over the last 12 months, how many different types of annual crops (incl. vegetables) did you cultivate?	Tick each crop cultivated out of a list of 19 annual crops.	Number of annual + perennial crops cultivated
	Over the last 12 months, how many different types of perennial crops (incl. vegetables) did you cultivate?	Tick each crop cultivated, out of a list of 8 perennial crops	
	And which of these annual crops (incl. vegetables) did you grow to sell on the market?	Tick each annual crop sold at market out of the list of crops cultivated. Total nr of annual crops sold (0-13)	Number of annual + perennial crops sold at market
	And which of these perennial crops (incl. vegetables) did you grow to sell on the market?	Tick each perennial crop sold at market out of the list of crops cultivated.	

Note that the resilience scale uses a simple count of the number of annual and perennial crops cultivated, and a simple count of the number of annual and perennial crops sold as its input.

Crop diversity is measured by counting the number of different crops farmers grow on their farms, and the number of different crops these households sell on the market. This captures both the diversity in what farmers grow and the diversity in what type of products farmers bring to the market.

Livestock situation

Table 9: Sub-construct Livestock Situation

Sub-construct	*Questionnaire item*	*Answer categories*	*Measures used*
Livestock situation	Which livestock does your household own?	Total nr of types of livestock owned (0-6)	Number of different type of livestock owned + number of different types of livestock products sold.
	Over the last 12 months, how many different livestock products did you sell on the market?	Total nr of types of livestock products sold (0-6)	

	Do you produce yourself sufficient feed on the farm to feed your own livestock?	Barely, no (1) – Yes (5)	Item response, score ranging from 1-5
	Are the fodder resources available during the dry season adequate for each kind of animal?	Barely, no (1) – Yes (5)	Item response, score ranging from 1-5.
Note that the resilience scale uses a simple count of the number of different types of livestock products owned and sold as its input. For the sufficiency of fodder and the availability of resources the (raw) item responses are used.			

The sub-construct of livestock situation is a combination of the 4 questionnaire items shown in table 9 above. Similar to crop diversity, we measure whether someone owns livestock and the extent to which farmers market livestock or livestock products. Originally, this sub-construct also included items that captured whether an animal's health was affected and the extent to which farmers could cope with dwindling health of their livestock. These items are excluded as they only captured a small portion of the sample. Merely a small share of farmers reported that the health of their livestock was negatively affected the past 12 months. Also note that in case the respondent does not own any livestock, these measures are set to 0 (zero).

Household resilience

Table 10: Sub-construct Household Resilience

Sub-construct	*Questionnaire item*	*Answer categories*	*Factor loading for item on HH-resilience*	*Measure used in composite hh-resilience scale*
Household resilience	Over the past year, rate the food situation of your household for each month. NB: average is taken over all twelve months.	Not enough (1) – Enough (5)	0.4579	Average on 8 item scale ranging from 1-low hh-resilience to -5- highest hh-resilience.
	How would you rate the health situation of your household in general?	Very bad (1) – Very good (5)	0.4861	
	Please explain whatever you know about Integrated Farm Management	No understanding (1) – Very good understanding (5)	0.6883	
	Does your household usually have sufficiently access to skilled persons (in or outside the household) to implement the plans you have for your household?	Not at all (1) – Yes, always (5)	0.6979	
	If there is a problem on your farm, does your household usually have the skills to solve it?	Idem	0.7248	
	Describe how in your household planning of tasks is done and divided among the members.	Not at all organised (1) – Very good organised (5)	0.7302	
	Describe how in your household decisions are taken on spending resources for farming inputs	Idem	0.7667	
	Describe how in your household decisions are taken on what crops to grow each season	Idem	0.7576	
Cronbach's α= 0.8554				

The eight questionnaire items in table 10 represent the sub-construct of household resilience. This household resilience measure mainly captures attittudes, and evidently a latent concept in the sense that we cannot directly observe a thing such as farmers' household resilience in such a way that we can directly observe by performing a direct count of the number of crops grown on a farm or the type of products a household sells on the market. Therefore we first investigated whether all these sub-

construct indeed have a single factorial structure so that they can be treated as a single concept. The factor analysis in Table 10 shows that the items do refer to a single latent construct and with tolerable factor loadings. The high cronbach's alpha demonstrates a high inter-item consistency. We use the average on the items in Table 10 as a measure of household resililience .

Household coping ability

Table 11: Sub-construct Coping Ability

Sub-construct	*Questionnaire item*	*Answer categories*	*Factor loading for item on HH-coping*	*Measure used in composite hh-coping scale*
Coping ability	Do you do anything different now compared to the past to cope with this first (main) shock?	Still unprepared (1) – Better prepared (5)	0.6268	Average on 8 item scale ranging from 1-low hh-coping to -5- highest hh-coping.
	Do you feel able now to cope with this first (main) shock if it would happen again?	Not at all able (1) – Very well able (5)	0.6629	
	Do you have all the assets and equipment you need to manage your farm as you would like to?	Not at all (1) – Yes absolutely (5)	0.7579	
	Do you have all the assets and equipment you need to manage your household as you would like to?	Idem	0.7293	
	Cronbach's α= 0.7679			

The sub-construct of coping ability is a combination of the four questionnaire items shown in the table above which captures to what extent farmers are able to deal with sudden and unforeseen incidents (exogenous shocks) that adversely affect their livelihood and financial situation. All respondents report having dealt with at least one shock over the past 5 years. The most frequently mentioned occurring shocks were death of a family member, serious illness within the family, or crop failure. See appendix 2.3, in the questionnaire, for a full overview. Whether respondents are able to mitigate the effects of such unfortunate events is again a latent construct that we cannot directly observe. A respondents' judgement on whether they possess sufficient assets and equipment to manage their farm and household is a similarly latent measure. Therefore we first checked whether the items are in fact referring to the same underlying latent construct through applying factor analyses. As can be seen in Table 11, each item has a factor loading above 0.6 which is indicative of high internal consistency across the items. Cronbach's alpha of 0.7679 indicates the constructed scale to be reliable.

Initially the scale included several extra items. First, it included "*Give the 3 main shocks that your household was exposed to in the last 5 years?*[14]" and "*How severe was the impact of the first (main) shock you mentioned on your household?*". The ability to recover from adverse events is an inherent part of one's resilience. The questionnaire inquired about a wide variety of adverse events outside the control of the respondents. Those mentioned, or listed, ranged from theft to death of family members, to crop-failure. As these shocks in themselves are of inherently different severity, referencing these shocks in follow up questions about their impacts, rendered this item incomparable across the households surveyed. In essence, we would be comparing, for instance, a theft or burgulary to a massive crop-failure or a fire on the farm. In such cases, the questions about the impact such a shock has less dependent on a household's adaptability or their ability of a household to recover but more a function of the severity of the shock experienced. The severity of the shock is largely outside the control of the household by definition, as we're focusing on exogenous shocks. Consequently, factor loadings for these items were below the threshold as well (respectively 0.2001 for experiencing shocks and

[14] Respondents were presented with a list of twelve shocks, where they could indicate the 1st, 2nd and 3rd main shock their household were exposed to. A new variable was created, where the main shock per household was categorized into 1) household shock, 2) economic shock, 3) climate shock.

0.2517 for severity). This demonstrates that these items are not consistent with the others, theoretically and empirically and were excluded from this sub-scale.

Note that the questionnaire also included items that asked for a subjective evaluation of the extent to which diversity of produce had changed the past 12 months. These measures are left out of these scales as we are comparing the generations of farmers (and the comparison group) at a set, static point in time (largely after the implementation of the programme). Including such measurements would already add a subjective evaluation of PAPAB's effectiveness (by the farmer) to the measurement of resilience, which would introduce noise in the resilience measurement (evidenced by the below-threshold values in the factor analysis as well). We did include analyses of these items under the heading other outcomes.

Resilience: Sub-construct to pillar

The overall resilience score takes in the measures on each sub-construct as described in the previous section. We have employed component analyses to allow for the combination of count-measures (the items referring to livestock situation and crop-diversity) and attitudinal measured. Finally, a resilience score is calculated as a weighted sum of the sub-constructs.

Table 12: Pillar 2 Resilience

Concept	*Sub-construct/ Component*	*Measure used in composite scale*	*Questionnaire items*	*Factor loading on resilience score*
Resilience	Crop diversity	Number of annual + perennial crops grown and number of annual and perennial crops sold.	Count of number of annual and perennial crops grown	0.6752
			Count of number of annual and perennial crops sold on market	0.7084
	Livestock situation	Number of different livestock owned and sold, having sufficient feed on the farm to feed own livestock, and availability of fodder during lean seasons (own fodder production and fodder bought).	Total livestock owned + Total number of livestock products sold	0.4394
			Do you produce yourself sufficient feed on the farm to feed your own livestock? (rescaled 1-5)	0.6226
			Are the fodder resources available during the dry season adequate for each kind of animal? (rescaled 1-5)	0.6362
	Household resilience	Household resilience: Attitudinal/subjective measure covering households capabilities of achieving household objectives and subjective assessments of intra-household decision making.	8 item scale, average between 1-5 see Table X for items	0.6154
	Household coping ability	Coping ability: Attitudinal/ subjective measure where respondents self-assess their ability to deal with sudden adverse events.	4 item scale average value 1-5, see Table X for items	0.7820

The concept of resilience is based on the 4 components of crop diversity, livestock situation, household resilience, and coping ability Table 12 above shows that each item within the components has a factor loading above the threshold, indicating internal consistency across items. Note that these factor loadings also serve as the weight for each sub-construct in the overall resilience score. The weighted sum was calculated and then normalized into a score between 0 – low resilience, and 100-high resilience to allow for a more meaningful interpretation and visualisation of the analyses related to resilience.

Pillar 3: Stewardship

Items to sub-constructs

Knowledge and awareness of changes in environment and sense of stewardship

Table 13: Sub-construct Knowledge & awareness of changes in environment & sense of stewardship

Sub-construct	*Questionnaire item*	*Answer categories*	*Factor loading*	*Measure used in composite scale*
Knowledge & awareness of changes in environment & sense of stewardship	Why do you think that the soil quality might change, what are possible reasons?	Not aware (1) – Very well aware (5)	0.7565	Average on 5 item scale ranging from 1-low awareness & sense of stewardship -5-highest awareness & sense of stewardship
	Why do you think that the vegetation might change, what are possible reasons?	Idem	0.7598	
	Why do you think that the water quality and quantity might change, what are possible reasons?	Not at all aware (1) – Very well aware (5)	0.6413	
	Please mention three concrete actions that you undertake to conserve protect natural resources outside your own farm.	Not at all clear response (1) – Very clear response (5)	0.5748	
	Please give ONE example that describes the importance of nature for yourself.	Idem	0.6269	
Cronbach's α= 0.8097				

The sub-construct of awareness of changes in the environment and a sense of stewardship combines the items presented in Table 13. Factor loadings are above threshold values, and the scale as has a high internal consistency. As can be seen, each item has a factor loading above 0.5, which indicates internal consistency across.

Initially, we expected the items to be able to combine the items "*Do you notice changes in the quality of the land or soil in your community?*", "*Do you notice changes in the vegetation in your community?*", and, *"Do you notice changes in the availability of water in your community?*" as well. However, these items have a lower bar of measuring awareness or knowledge about changes in the environment as saying one notices changes is different than knowing *why,* for instance, the water quality changes. Also, we find a smaller variance on the items asking whether people notice changes. This indicates that most say they notice changes, but when pressed and asked about such changes further (e.g. '*why* do you think water quality might change'), thus giving more substance to that awareness the variance in responses is bigger. Empirically, we found that the items that ask whether people notice changes did not meet the threshold for inclusion in a scale when performing factor analyses.

Additionally, find that the items that ask for examples of actions that people undertake to protect natural resources, and the question to describe the importance of nature are strongly associated with the 'why' items about soil, vegetation, and, water quality. Initially, we envisioned the items on actions to protect natural resources outside the farm, and the importance people attach to nature to be a separate construct (sense of stewardship), but that did not hold empirically. Based on the associations between

the items, we cannot distinguish between The item, "*Do you consider it important to protect the environment and its natural resources?*" is missing because that item did not load appropriately on the scale generated. We suspect that the reason for this is similar as for the items referring to noticing change; one can gratuitously state that something is important, but when pressed to give an example ('Please give ONE example that describes the importance of nature for yourself') we see that the items are more discriminatory and thus better able to distinguish the sense of stewardship that respondents have.

Knowledge and awareness on use of the commons

Table 14: Knowledge & awareness on use of the commons

Sub-construct	*Questionnaire item*	*Factor loading*	*Answer categories*	*Measure used in composite scale*
Knowledge & awareness on use of the commons	Please explain about the importance of trees and bushes on the land outside your own farm.	0.5303	Not at all aware (1) – Very well aware (5)	Average on 6 item scale ranging from 1-low conservation of commons -5- high conservation of commons
	Please explain how you use trees and bushes on the land outside your own farm.	0.5775	Not good, exploiting (1) – Very well, conserving (5)	
	Please explain the importance of water sources in this village and how you use them.	0.6271	Idem	
	Please describe – with examples – how you try to conserve water yourself?	0.6140	Idem	
	Please explain the importance of the common lands in this village and how you use them.	0.5943	Idem	
	Please describe – with examples – how you try to conserve the common lands yourself?	0.6005	Idem	
Cronbach's α= 0.7760				

The six questionnaire items in the table above together measure the sub-construct of use of the commons. All factor loadings are above 0.5, indicating internal consistency across items. Cronbach's alpha of 0.7488 indicates the constructed scale for the use of the commons to be reliable. This structure of this scale is consistent with that envisioned during the questionnaire design stage.

Knowledge and awareness about farming practices

Table 15: Knowledge and awareness about farming practices

Sub-construct	*Questionnaire item*	*Factor loading*	*Answer categories*	*Measure used in composite scale*
Knowledge and awareness about farming practices	If you have such practices [physical practices], please explain why and how you apply these practices.	0.6947	Not at all aware (1) – Very well aware (5)	Average on 5 item scale ranging from 1-low knowledge & awareness of farming practices to -5- high knowledge and awareness of farming practices
	If you have such practices [land management practices], please explain **why and how** you apply these practices.	0.7099	Not at all aware (1) – Very well aware (5)	
	Please describe if you have trees on your farm, how you use them, why you have them and if you are satisfied?	0.5645	Not at all aware (1) – Very well aware (5)	
	Why do you use crop rotations on your fields?	0.7810	Not at all aware (1) – Very well aware (5)	
	If you have such practices [fertilizer use], please explain **why and how** you apply these practices.	0.7535	Not at all aware (1) – Very well aware (5)	

Cronbach's α= 0. 0.8425	

In our conceptualisation of stewardship, we differentiate between knowledge and actual behaviour. Knowledge is a softer, more latent-type measure, whereas practices can be directly observed. In this case, we thus distinguish the actual implementation of practices (see below) from knowing why and how to apply these practices (see Table 15). The sub-construct of knowledge and awareness about farming practices is asked by having farmers explain what they know about several practices. Subsequently, the interviewer (who is knowledgeable about such practices) judges the adequacy of the response. The sub-construct knowledge and awareness about farming practices thus measures the extent to which respondents are knowledgeable about a variety of practices; that is, knowing when, how, and why to implement the given set of practices. We find that all knowledge-related questions conform to a single factorial structure appropriately high factor loadings and high internal consistency (Cronbach's alpha > 0.8).

Farming practices

The final components of the sub-constructs stewardship measures whether farmers actually implement a set of farming practices. These are physical practices, land management practices, crop rotation and soil management practices (mainly related to the use of fertilizer). We have included all practices measured in the survey. All practices are analysed separately, meaning that we assess whether the proportion of respondents implementing a practice differs between the target and comparison group.

Table 16 Farming practices

Questionnaire item	*Questionnaire item*	*Answer categories*	*Measure used in composite sale*
Physical practices	**Physical**: Implements Contourlines/trenches, Billon continu, Stone bunds, Gully control	Yes/No (for each practice)	Analyses on each practice separately (proportion of people who implement a practice)
Land management practices	**Land management:** Ploughing on the contourline, Quinconce / staggered row planting, Mulching, Cover crops	Yes/No (for each practice)	Analyses on each practice separately (proportion of people who implement a practice)
Crop rotation	**Crop rotation**: Implements crop -rotation on most or all plots of land	Implements crop rotation on most or all plots of land (responses for crop rotation on none or some plots of land coded as 0).	Analyses on each practice separately (proportion of people who implement crop rotation on most plots of land)
Soil management practices	**Soil**: Uses compost, manure, chemical fertilizer, combine compost+chemical fertilizer, combine manure + chemical fertilizer	Yes/No (for each practice)	Analyses on each practice separately (proportion of people who implement a practice)

Stewardship: sub-constructs to pillar

The composite stewardship score is a combination of as the sub-constructs are a combination of latent (knowledge and awareness) and directly observable items (practices). We combine all these measures into a single composite score from 0 – lowest stewardship to -100- highest level of stewardship. The stewardship score is a weighted sum on the scores on each sub-construct, where the weights are determined by the factor loadings presented in Table 17. These sum-scores are then re-scaled to range between 0 and 100 for ease of interpretation. We cannot use each farming practice separately the overall stewardship score (Note that we do, however present analyses of each practice separately in section 3.2.3. Given its highly skewed distributions, i.e. there are only a very small share of respondents that for instance implement gully control, this would lead to a disproportionally heavy weight being placed on such individual practices in the overall stewardship score. Therefore, we used a count, that is the number of different physical, land management, and soil management practice a respondent implements, as input to the overall stewardship scale. Additionally, note that generally, the knowledge and awareness items carry the most weight in the construction of overall pillar score.

Table 17: Pillar 3 Stewardship

Concept	*Sub-construct/ Component*	*Measurement*	*Questionnaire items*	*Weight/component loading on stewardship scale*
Stewardship	Knowledge & awareness of changes in environment & sense of stewardship	Awareness: Subjective assessment of respondents' awareness on soil, vegetation, and water quality.	5 item scale, average between 1-5 see Table 13 for items	0.4502
	Knowledge & awareness on use of the commons	Use of the commons: Subjective assessment of respondents' awareness on conservation of commons	Average on 6 item scale ranging from 1- low conservation of commons -5- high conservation of commons	0.4360
	Knowledge and awareness about farming practices/ how and why to implement practices	Knowledge and awareness about farming, land management and soil practices/ How and why to implement such practices.	Average on 5 item scale ranging from 1- low knowledge & awareness of farming practices to -5- high knowledge and awareness of farming practices	0.4806
	Physical practices	Count of physical practices implemented.	Physical practices yes/no see Table 16	0.3074
	Land management practices	Count of land management practices implemented.	Farm management practices yes/no see Table 16	0.3806
	Crop rotation	Uses crop rotation on most plots	Uses crop rotation on most plots, see Table 16.	0.2580
	Soil management practices	Count of soil management practices implemented.	Soil management practices yes/no see Table 16	0.2651

2.2 PROPENSITY SCORE MATCHING: 4 TARGET GROUPS, 4 COMPARISON GROUPS

		PIP Generation 1						PIP Generation 2						PIP Generation 3						PIP Generation 4						All PIP					
		Pre			*Post*			*Pre*			*Post*			*Pre*			*Post*			*Pre*			*Post*			*Pre*			*Post*		
		PIP G1	C	PIP G1-C	PIP G1	C	PIP G1-C	PIP G2	C	PIP G2-C	PIP G2	C	PIP G2-C	PIP G3	C	PIP G3-C	PIP G3	C	PIP G3-C	PIP G4	C	PIP G4-C	PIP G4	C	PIP G4-C	All PIP	C	All PIP-C	All PIP	C	All PIP-C
		Mean/[SE]		Δt-test	Mean/[SE]		Δt-test	Mean/[SE]		Δt-test	Mean/[SE]		Δt-test	Mean/[SE]		Δt-test	Mean/[SE]		Δt-test	Mean/[SE]		Δt-test	Mean/[SE]		Δt-test	Mean/[SE]		Δt-test	Mean/[SE]		Δt-test
Province	Bujumbura (%)	0.113	0.255	-0.142***	0.247	0.250	-0.003	0.261	0.255	0.006	0.336	0.349	-0.012	0.100	0.255	-0.155***	0.217	0.197	0.019	0.300	0.255	0.045	0.237	0.267	-0.030	0.200	0.255	-0.055**	0.224	0.257	-0.033
		[0.031]	[0.020]		[0.063]	[0.024]		[0.040]	[0.020]		[0.065]	[0.026]		[0.028]	[0.020]		[0.054]	[0.019]		[0.039]	[0.020]		[0.040]	[0.021]		[0.018]	[0.020]		[0.023]	[0.021]	
	Cibitoke (%)	0.236	0.255	-0.019	0.225	0.213	0.012	0.000	0.255	-0.255***	0.000	0.000	N/A	0.392	0.255	0.137***	0.334	0.280	0.054	0.100	0.255	-0.155***	0.211	0.224	-0.014	0.177	0.255	-0.077***	0.288	0.229	0.059
		[0.041]	[0.020]		[0.055]	[0.023]		[0.000]	[0.020]		[0.000]	[0.000]		[0.045]	[0.020]		[0.051]	[0.022]		[0.025]	[0.020]		[0.051]	[0.020]		[0.017]	[0.020]		[0.033]	[0.021]	
	Makamba (%)	0.160	0.121	0.039	0.107	0.110	-0.003	0.361	0.121	0.240***	0.144	0.164	-0.020	0.108	0.121	-0.013	0.093	0.121	-0.028	0.200	0.121	0.079**	0.100	0.122	-0.023	0.208	0.121	0.087***	0.135	0.124	0.011
		[0.036]	[0.015]		[0.034]	[0.017]		[0.044]	[0.015]		[0.031]	[0.020]		[0.028]	[0.015]		[0.031]	[0.016]		[0.034]	[0.015]		[0.024]	[0.015]		[0.018]	[0.015]		[0.018]	[0.016]	
	Muyinga (%)	0.245	0.127	0.118***	0.123	0.159	-0.035	0.252	0.127	0.125***	0.216	0.176	0.040	0.200	0.127	0.073**	0.132	0.143	-0.010	0.200	0.127	0.073**	0.145	0.133	0.012	0.223	0.127	0.095***	0.121	0.143	-0.022
		[0.042]	[0.015]		[0.042]	[0.020]		[0.040]	[0.015]		[0.050]	[0.021]		[0.037]	[0.015]		[0.032]	[0.017]		[0.034]	[0.015]		[0.033]	[0.016]		[0.019]	[0.015]		[0.015]	[0.017]	
	Rumonge (%)	0.245	0.242	0.003	0.297	0.268	0.029	0.126	0.242	-0.116***	0.304	0.311	-0.007	0.200	0.242	-0.042	0.224	0.259	-0.035	0.200	0.242	-0.042	0.308	0.253	0.055	0.192	0.242	-0.050*	0.232	0.248	-0.016
		[0.042]	[0.020]		[0.058]	[0.025]		[0.031]	[0.020]		[0.070]	[0.025]		[0.037]	[0.020]		[0.048]	[0.021]		[0.034]	[0.020]		[0.053]	[0.021]		[0.018]	[0.020]		[0.025]	[0.021]	
Gender	Female (%)	0.509	0.454	0.055	0.402	0.448	-0.046	0.462	0.454	0.008	0.412	0.434	-0.022	0.633	0.454	0.179***	0.504	0.480	0.024	0.564	0.454	0.110**	0.403	0.456	-0.052	0.544	0.454	0.090***	0.458	0.448	0.010
		[0.049]	[0.023]		[0.064]	[0.028]		[0.046]	[0.023]		[0.067]	[0.027]		[0.044]	[0.023]		[0.057]	[0.024]		[0.042]	[0.023]		[0.051]	[0.024]		[0.023]	[0.023]		[0.030]	[0.024]	
Age	Age (mean)	44.415	41.841	2.574*	43.987	42.595	1.392	42.874	41.841	1.033	42.161	43.290	-1.129	39.367	41.841	-2.474*	41.260	41.173	0.086	40.657	41.841	-1.184	40.881	41.760	-0.879	41.703	41.841	-0.138	42.315	41.245	1.070
		[1.173]	[0.646]		[1.802]	[0.759]		[1.189]	[0.646]		[1.869]	[0.769]		[1.340]	[0.646]		[1.679]	[0.653]		[0.965]	[0.646]		[1.224]	[0.662]		[0.586]	[0.646]		[0.746]	[0.658]	
Education	No education (%)	0.066	0.210	-0.144***	0.107	0.119	-0.012	0.227	0.210	0.017	0.155	0.208	-0.053	0.225	0.210	0.015	0.232	0.209	0.023	0.150	0.210	-0.060	0.179	0.196	-0.017	0.169	0.210	-0.041	0.209	0.181	0.028
		[0.024]	[0.019]		[0.039]	[0.018]		[0.039]	[0.019]		[0.040]	[0.022]		[0.038]	[0.019]		[0.048]	[0.020]		[0.030]	[0.019]		[0.045]	[0.019]		[0.017]	[0.019]		[0.030]	[0.019]	
	Ecole d'alphabetisation	0.132	0.221	-0.089**	0.195	0.216	-0.021	0.235	0.221	0.014	0.233	0.214	0.019	0.267	0.221	0.046	0.299	0.228	0.071	0.243	0.221	0.022	0.222	0.220	0.002	0.223	0.221	0.002	0.232	0.224	0.008
		[0.033]	[0.019]		[0.053]	[0.023]		[0.039]	[0.019]		[0.054]	[0.022]		[0.041]	[0.019]		[0.054]	[0.020]		[0.036]	[0.019]		[0.044]	[0.020]		[0.019]	[0.019]		[0.026]	[0.020]	
	Ecole primaire (%)	0.642	0.444	0.198***	0.516	0.509	0.007	0.471	0.444	0.027	0.474	0.455	0.020	0.433	0.444	-0.010	0.376	0.458	-0.082	0.486	0.444	0.042	0.463	0.460	0.003	0.503	0.444	0.059*	0.456	0.467	-0.011
		[0.047]	[0.023]		[0.066]	[0.028]		[0.046]	[0.023]		[0.068]	[0.027]		[0.045]	[0.023]		[0.053]	[0.024]		[0.042]	[0.023]		[0.054]	[0.024]		[0.023]	[0.023]		[0.030]	[0.024]	
	College lycee or higher (%)	0.160	0.125	0.035	0.182	0.155	0.026	0.067	0.125	-0.058*	0.137	0.123	0.014	0.075	0.125	-0.050	0.092	0.105	-0.012	0.121	0.125	-0.004	0.136	0.124	0.012	0.105	0.125	-0.020	0.103	0.129	-0.025
		[0.036]	[0.015]		[0.055]	[0.020]		[0.023]	[0.015]		[0.061]	[0.018]		[0.024]	[0.015]		[0.037]	[0.015]		[0.028]	[0.015]		[0.040]	[0.016]		[0.014]	[0.015]		[0.016]	[0.016]	
Household head type	Dual headed (%)	0.858	0.728	0.130***	0.840	0.780	0.059	0.790	0.728	0.062	0.681	0.757	-0.075	0.750	0.728	0.022	0.762	0.743	0.018	0.786	0.728	0.057	0.708	0.738	-0.030	0.794	0.728	0.066**	0.780	0.745	0.035
		[0.034]	[0.021]		[0.044]	[0.023]		[0.038]	[0.021]		[0.068]	[0.023]		[0.040]	[0.021]		[0.046]	[0.021]		[0.035]	[0.021]		[0.052]	[0.021]		[0.018]	[0.021]		[0.025]	[0.021]	
	Female headed (%)	0.038	0.127	-0.090***	0.054	0.079	-0.025	0.109	0.127	-0.018	0.110	0.120	-0.010	0.150	0.127	0.023	0.141	0.133	0.008	0.086	0.127	-0.042	0.093	0.118	-0.024	0.097	0.127	-0.030	0.115	0.112	0.003
		[0.019]	[0.015]		[0.028]	[0.015]		[0.029]	[0.015]		[0.040]	[0.018]		[0.033]	[0.015]		[0.038]	[0.017]		[0.024]	[0.015]		[0.031]	[0.015]		[0.013]	[0.015]		[0.020]	[0.015]	
	Male headed (%)	0.104	0.144	-0.041	0.106	0.140	-0.035	0.101	0.144	-0.044	0.208	0.123	0.085	0.100	0.144	-0.044	0.097	0.124	-0.026	0.129	0.144	-0.016	0.199	0.144	0.054	0.109	0.144	-0.035	0.105	0.143	-0.038
		[0.030]	[0.016]		[0.037]	[0.019]		[0.028]	[0.016]		[0.065]	[0.018]		[0.028]	[0.016]		[0.031]	[0.016]		[0.028]	[0.016]		[0.047]	[0.017]		[0.014]	[0.016]		[0.017]	[0.017]	
Farm type	Mixed farm (%)	0.953	0.605	0.348***	0.854	0.832	0.022	0.857	0.605	0.252***	0.732	0.680	0.051	0.758	0.605	0.153***	0.603	0.639	-0.036	0.836	0.605	0.231***	0.664	0.629	0.035	0.847	0.605	0.242***	0.683	0.674	0.009
		[0.021]	[0.023]		[0.059]	[0.021]		[0.032]	[0.023]		[0.065]	[0.025]		[0.039]	[0.023]		[0.057]	[0.023]		[0.031]	[0.023]		[0.057]	[0.023]		[0.016]	[0.023]		[0.035]	[0.023]	
Land hectares	Land hectares (ha)	2.447	1.593	0.854***	1.871	1.854	0.017	2.044	1.593	0.451***	1.875	1.745	0.130	1.497	1.593	-0.096	1.511	1.560	-0.050	1.863	1.593	0.270**	1.672	1.613	0.059	1.945	1.593	0.351***	1.728	1.675	0.052
		[0.154]	[0.067]		[0.143]	[0.089]		[0.117]	[0.067]		[0.137]	[0.077]		[0.079]	[0.067]		[0.106]	[0.071]		[0.103]	[0.067]		[0.105]	[0.067]		[0.059]	[0.067]		[0.063]	[0.071]	
Land ownership	Land ownership (%)	0.911	0.841	0.071***	0.869	0.881	-0.012	0.877	0.841	0.036	0.883	0.883	0.000	0.809	0.841	-0.031	0.810	0.827	-0.017	0.865	0.841	0.025	0.869	0.847	0.023	0.864	0.841	0.024	0.846	0.855	-0.009
		[0.016]	[0.012]		[0.023]	[0.012]		[0.016]	[0.012]		[0.026]	[0.012]		[0.022]	[0.012]		[0.026]	[0.013]		[0.017]	[0.012]		[0.019]	[0.012]		[0.009]	[0.012]		[0.013]	[0.012]	
N		106	471		105	328		119	471		119	341		120	471		120	421		140	471		140	450		485	471		483	420	

The value displayed for t-tests are the differences in the means across the groups.
***, **, and * indicate significance at the 1, 5, and 10 percent critical level.

Figure 33 Propensity score matching, pre- and post-matching balance test for characteristics included in matching.

We use the propensity score matching to solve the problem of incomparability between the target and comparison group in two stages. In the first stage, we calculate the propensity score in order to select or match a comparison group where the distribution of the aforementioned socio-demographics, farm characteristics, and location are similar to those in the target group. Finding these matches is done based on the propensity scores calculated. Each person in the comparison group receives a weight, based on their propensity score. This weight can colloquially be interpreted as a measure of similarity between that particular person in the comparison group and its match in the target group. Second, we calculate the values on the relevant outcome indicator for the comparison group using a weight for each observation in the comparison group. By doing so, bad matches, or in other words, people that are not very comparable to those in the target group, receive a lower weight in the calculation of the outcome for the counterfactual (comparison group). Better matches, or people in the comparison group who are more comparable to the people in the target group, receive a higher weight. By doing so, we make sure that the target and comparison group are comparable and balanced while still employing a large share of the sample that we have collected

We have used a normal (Gaussian) kernel estimator, where each person in the comparison is given a weight based on the characteristics used in the matching model. his weight is a kernel-weighted average of the distance between a given person in the target group to all people in the comparison group, where the weight is expressed as a proportion of closeness between a subject in the comparison group and the target group. Subsequently, when calculating the average values on the outcome indicator for people in the comparison group, each person in the comparison group is given a weight, so that closer and better matches, thus more comparable people, have a greater influence on this average compared to worse matches.

2.3 MULTIVARIATE ANALYSES: ASSESSING RELATIONSHIPS BETWEEN PILLARS.

	Full sample		All PIP		PIP Generation 1		PIP Generation 2		PIP Generation 3		PIP Generation 4		Comparison group	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
	Stewardship	Resilience	Stewardship	Resilience	Stewardship	Resilience	Stewardship	Resilience	Stewardship	Resilience	Stewardship	Resilience	Stewardship	Resilience
Motivation	0.43087***		0.25026***		0.18008		0.21556		0.05535		0.33746**		0.35246***	
	(0.000)		(0.008)		(0.351)		(0.128)		(0.632)		(0.049)		(0.000)	
Stewardship		0.33066***		0.15951**		-0.40458**		0.02076		-0.05326		0.23471**		0.24837***
		(0.000)		(0.032)		(0.034)		(0.859)		(0.675)		(0.018)		(0.000)
Covariates	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Constant	34.73840***	6.31817	52.40877***	29.75214***	75.42074***	69.07486***	52.76069***	49.65367***	64.21767***	36.34462***	50.46121***	25.42053*	39.78116***	7.41314
	(0.000)	(0.170)	(0.000)	(0.000)	(0.000)	(0.005)	(0.002)	(0.006)	(0.000)	(0.000)	(0.001)	(0.056)	(0.000)	(0.148)
Observations	950	946	486	484	107	107	119	118	120	120	140	139	464	462
R-squared	0.386	0.526	0.211	0.438	0.255	0.432	0.237	0.395	0.180	0.512	0.285	0.471	0.288	0.511
Covariates included: province, gender, age, education, type of household head, farm type, # hectares, % of land that is owned														
Robust pval in parentheses														
*** p<0.01, ** p<0.05, * p<0.1														

Figure 34 Results of multivariate regression analyses (OLS) for relationships between pillars for full sample (weighted) and disaggregated by generation

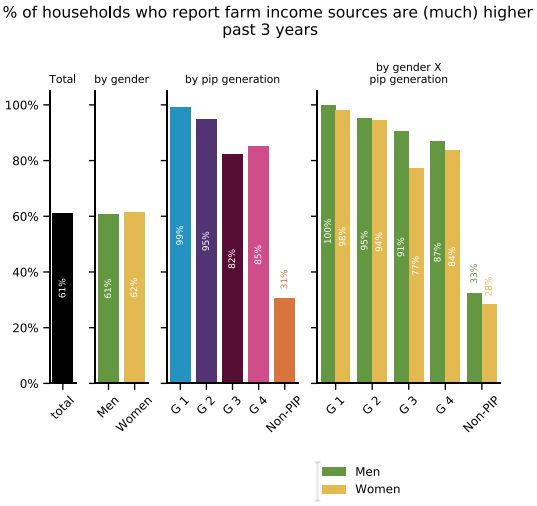
2.4 QUESTIONNAIRE

Please follow this link to the full questionnaire.

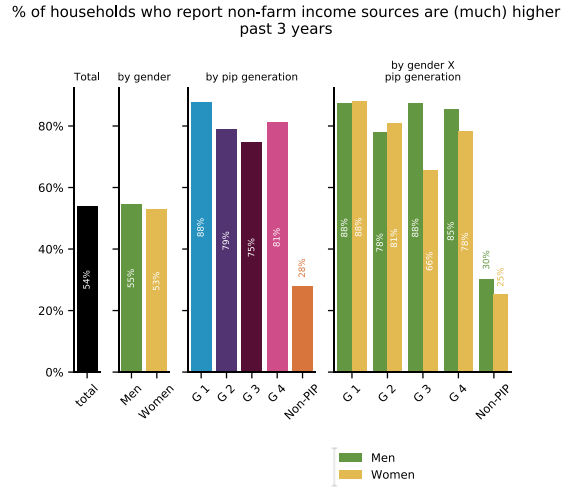
2.5 INDICATORS FROM PIP-IMPACT STUDY TO BE USED FOR THE FINAL EVALUATION OF THE PAPAB-PROJECT.

Some of the data collected in the PIP-impact study is to be used as input to the final evaluation. The following graphs present these indicators relevant to the final evaluation. These following graphs present the total of the whole sample, the outcomes disaggregated by gender, disaggregated by PIP-farmer generation, and disaggregation by gender and farmers generation (e.g. women in generation 1 versus men in generation 1).

Outcome type:	Project impact indicators
Outcome:	GL 1.a: The proportion of households that report an improvement in average annual income. (farm and non-farm income)



Source: PIP-impact study, n=941



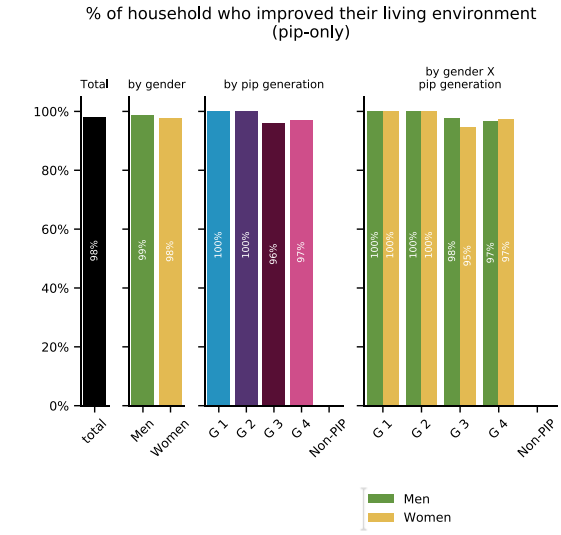
Source: PIP-impact study, n=527

Percentage of households who report farm income sources are higher or much higher the past 3 years

Group		*%*
Total		61%
Men		61%
Women		62%
G 1		99%
G 2		95%
G 3		82%
G 4		85%
Non-PIP		31%
G 1	Men	100%
G 1	Women	98%
G 2	Men	95%
G 2	Women	94%
G 3	Men	91%
G 3	Women	77%
G 4	Men	87%
G 4	Women	84%
Non-PIP	Men	33%
Non-PIP	Women	28%

Percentage of households who report non-farm income sources are higher or much higher the past 3 years	
Group	*%*
Total	54%
Men	55%
Women	53%
G 1	88%
G 2	79%
G 3	75%
G 4	81%
Non-PIP	28%
G 1 Men	88%
G 1 Women	88%
G 2 Men	78%
G 2 Women	81%
G 3 Men	88%
G 3 Women	66%
G 4 Men	85%
G 4 Women	78%
Non-PIP Men	30%
Non-PIP Women	25%

Outcome type:	Project impact indicators
Outcome:	GL 1.b: The proportion of households that report having improved their living environment. (pip only)



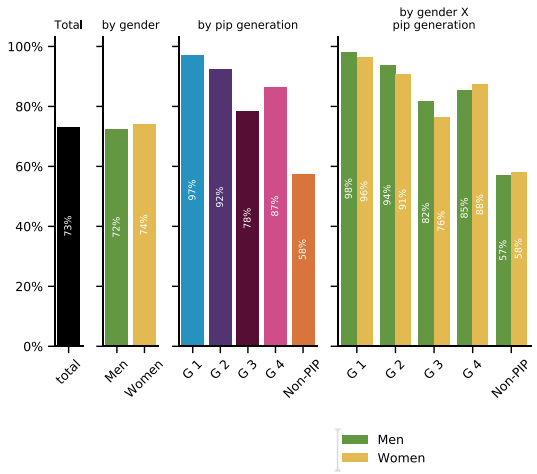
Source: PIP-impact study, n=488

Percentage of households who improved their living environment (pip-only)

Group		*%*
Total		*98%*
Men		99%
Women		98%
G 1		100%
G 2		100%
G 3		96%
G 4		97%
Non-PIP		-
G 1	Men	100%
G 1	Women	100%
G 2	Men	100%
G 2	Women	100%
G 3	Men	98%
G 3	Women	95%
G 4	Men	97%
G 4	Women	97%
Non-PIP	Men	-
Non-PIP	Women	-

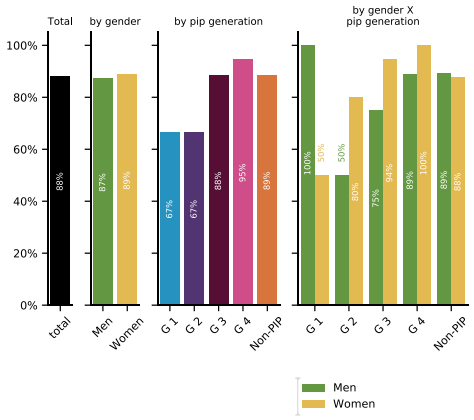
Outcome type:	2nd component objectives:
Outcome:	**SO 23: Proportion of Beneficiaries in the intervention zones declaring having an added value (interest) in associating as a group.**

% of households who are associated in a group



Source: PIP-impact study, n=962

% of non-associated households who are interested in associating in a group



Source: PIP-impact study, n=258

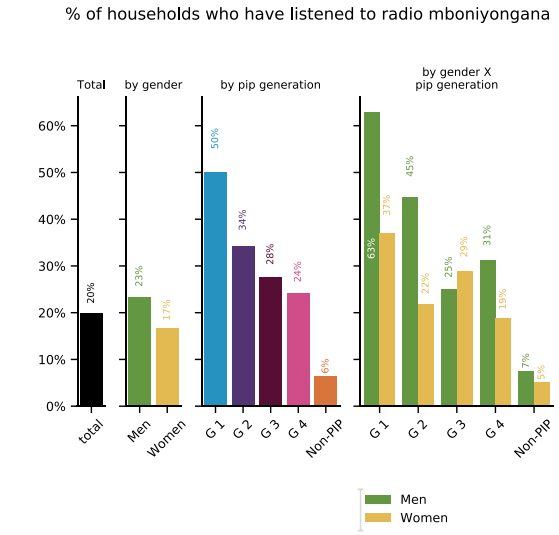
Note that 73% of respondents reports to that they are already associated in a group, 27% (n=258) of respondents reports that they are not associated with a group, yet.

Percetnage of households who are associated in a group

Group		*%*
Total		73%
Men		72%
Women		74%
G 1		97%
G 2		93%
G 3		78%
G 4		87%
Non-PIP		58%
G 1	Men	98%
G 1	Women	96%
G 2	Men	94%
G 2	Women	91%
G 3	Men	82%
G 3	Women	76%
G 4	Men	85%
G 4	Women	88%
Non-PIP	Men	57%
Non-PIP	Women	58%

Percentage of non-associated households who are interested in associating in a group		
	Group	*%*
	Total	88%
	Men	87%
	Women	89%
	G 1	67%
	G 2	67%
	G 3	88%
	G 4	95%
	Non-PIP	89%
G 1	Men	100%
G 1	Women	50%
G 2	Men	50%
G 2	Women	80%
G 3	Men	75%
G 3	Women	94%
G 4	Men	89%
G 4	Women	100%
Non-PIP	Men	89%
Non-PIP	Women	88%

Outcome type:	Indicators of outcome 2.3 :
Outcome:	Households who listened to radio messages (radio Mboniyongana)



Source: PIP-impact study, n=962

Percentage of households who have listened to radio Mboniyongana

Group		*%*
Total		20%
Men		23%
Women		17%
G 1		50%
G 2		34%
G 3		28%
G 4		24%
Non-PIP		6%
G 1	Men	63%
G 1	Women	37%
G 2	Men	45%
G 2	Women	22%
G 3	Men	25%
G 3	Women	29%
G 4	Men	31%
G 4	Women	19%
Non-PIP	Men	7%
Non-PIP	Women	5%

This report has been produced by the Impact Measurement and Knowledge team and is written by Rik Linssen and Marieke Meeske

For more information, or to comment on this publication, please email:
rik.linssen@oxfamnovib.nl

Published by Oxfam Novib in July 2020.

Oxfam Novib
P.O. Box 30919
2500 GX The Hague
The Netherlands

T +31 (0) 70 3421621
info@oxfamnovib.nl
www.oxfamnovib.nl