Journal of Rural Studies 98 (2023) 34–58

Contents lists available at ScienceDirect

Journal of Rural Studies

journal homepage: www.elsevier.com/locate/jrurstud

Do organic farming initiatives in Sub-Saharan Africa improve the sustainability of smallholder farmers? Evidence from five case studies in Ghana and Kenya

Johan Blockeel[a,*], Christian Schader[a], Anja Heidenreich[a], Christian Grovermann[a], Irene Kadzere[a], Irene S. Egyir[b], Anne Muriuki[c], Joseph Bandanaa[b], Chrysantus M. Tanga[d], Joseph Clottey[b], John Ndungu[c], Matthias Stolze[a]

[a] Research Institute of Organic Agriculture, Ackerstrasse 113, Frick, Switzerland
[b] University of Ghana, Legon, Accra, Ghana
[c] Kenya Agricultural and Livestock Research Organization, Box 220-01000, Thika, Kenya
[d] International Centre of Insect Physiology and Ecology, P.O. Box 30772-00100, Nairobi, Kenya

ARTICLE INFO

Keywords:
Sustainability assessment
Impact evaluation
Smallholder farming systems
Organic agriculture

ABSTRACT

Organic agriculture (OA) is often regarded as a sustainable agricultural pathway for smallholder farmers in Sub-Saharan Africa, and an increasing number of initiatives promoting OA were initiated over the last decades. However, holistic empirical evidence on the effects of such initiatives on the sustainability of smallholder farmers is still scanty. We analyzed the effects of five initiatives promoting OA on farm-level sustainability. We selected farmers exposed to the initiatives (n = 678) and control farms (n = 957) in five different case studies, two implemented in Ghana and three in Kenya. We used a farm-level multi-criteria assessment tool that evaluates to what extent the environmental, social, economic, and governance sustainability goals formulated in the FAO-SAFA Guidelines are addressed by farmers. We found that the initiatives had limited effects on reducing farmers reliance on chemical inputs use (pesticides and synthetic fertilizers) and uptake of organic or agro-ecological practices. Nevertheless, the results show that the initiatives were able to trigger significant (*p-value* < 0.05) positive effects mainly for the environmental sustainability goals. In contrast, the goals within the economic, social and good governance sustainability dimensions were rarely affected. Moreover, certified initiatives had more frequently a positive sustainability effect compared to uncertified initiatives.

1. Introduction

Several sustainable agricultural pathways emerged to tackle the complex set of interrelated environmental (e.g. land degradation, climate change) and socio-economic challenges (e.g. food security, price volatility) linked with agricultural production (Oberč, 2020). Among these, organic agriculture (OA) is today one of the most widely spread alternatives globally (Willer and Lernoud, 2019). Particularly for marginalized smallholder farmers in developing countries – lacking access to resources, information, technology, capital, and assets (IFAD, 2013) - it is considered an opportunity to improve livelihoods and protect the environment (Jouzi et al., 2017). In Sub-Saharan Africa (SSA) – where agriculture is mainly dominated by small-scale farms (NEPAD, 2013) - various stakeholders recognized that OA can play a significant role in the achievement of the Sustainable Development Goals as well as the African Unions Agenda 2063 (UNEP-UNCTAD, 2008).

In the Codex Alimentarius guidelines, OA is defined as "*a holistic production management [whose] primary goal is to optimize the health and productivity of interdependent communities of soil, life, plants animals and people*" (FAO and WHO, 2007, p2). In the African context, different types of OA can be distinguished: certified OA, uncertified OA, and "organic by default" agriculture (Parrot et al., 2003; Twarog, 2006). Certified OA has a dedicated regulatory framework and requires third-party certification to attest conformity with production standards. Since individual certification would be unaffordable and administratively too complex to

* Corresponding author. Department of Socio-Economic Sciences, Research Institute of Organic Agriculture FiBL, Ackerstrasse 113, Postfach 219, 5070, Frick, Switzerland.

E-mail address: johan.blockeel@fibl.org (J. Blockeel).

https://doi.org/10.1016/j.jrurstud.2023.01.010
Received 26 June 2022; Received in revised form 12 January 2023; Accepted 17 January 2023
Available online 7 February 2023

manage for most smallholders, a system of "group-certification" was introduced (Meinshausen et al., 2019). Driven by the increasing demand for certified commodities from the global North (Anderberg, 2020; Parrott et al., 2006; Rosinger, 2013), the area under certified OA increased by 10,2% between 2018 and 2019 and reached 2,8% of the continent's agricultural area in 2019 (Willer and Lernoud, 2019). Non-certified OA differs from this type as it is not subject to inspection, certification, and labeling (Scialabba, 2002). The latter is considered widespread over the continent (Parrott et al., 2006; Rosinger, 2013), but exact figures were not available. The term is often associated, but not to be confused, with "traditional" or "organic by default" farming (Twarog, 2006). This type is described as a low-input low-output system (Kamau et al., 2018), mainly occurring due to the inability of smallholders to access or afford synthetic fertilizers and plant protection products (Bennett and Franzel, 2013).

Both certified and uncertified OA are commonly promoted among smallholders by external actors, either private companies aiming to export the produce or national and international development agencies aiming at improving farmers' livelihoods (Parrot et al., 2003). Such initiatives are diverse in their intervention methods and outcomes. Still, they commonly rely on capacity development (e.g. group training, farmer field schools, or demonstration plots) to promote the uptake of OA (ITC, 2016). The latter is an important activity given the knowledge-intensive nature of OA (Tripp, 2001) and is expected to trigger changes in management practices that provide economic, social, and environmental benefits (Oya et al., 2017). On the other hand, a market intervention component (i.e. premium prices) is usually included in certified initiatives. The latter is expected to have more direct positive socio-economic effects. Nevertheless, the expected outcomes of the market and capacity development components largely overlap (ITC, 2016; Oya et al., 2017).

However, holistic and credible evidence on the sustainability effects of different types of initiatives promoting OA (i.e. certified and non-certified) within the smallholder African context is still scarce (Bennett and Franzel, 2013; De Fries et al., 2017; Oya et al., 2017). Previous studies done in African countries mainly investigated economic effects (e.g. yields, premium prices, net income) of certified OA initiatives or other certification schemes (Bolwig et al., 2009; Chiputwa et al., 2015; Gibbon et al., 2009; Jena et al., 2012; Meemken, 2020; Ruben, 2015), thus did not account for the multidimensional impact pathways associated with OA (Kilcher, 2007). Among the studies taking a holistic approach, Ssebunya et al., 2019 analyzed the sustainability of certified organic and Fair Trade coffee farmers in Uganda using the Sustainability Monitoring and Assessment RouTine-Farm Tool (SMART-Farm Tool) (Schader et al., 2016). Nevertheless, no rigorous methods controlling for the risk of selection bias were applied, a commonly found methodological shortcoming in evaluation studies of sustainability standards or agricultural interventions (De Fries et al., 2017; Oya et al., 2017; Stewart et al., 2015). Additionally, past studies did not capture the diversity of settings in which OA is implemented. In particular, non-certified initiatives promoting OA remained largely understudied in past studies (Ansah et al., 2020; Bennett and Franzel, 2013; Oya et al., 2017) even though, based on anecdotal evidence they are expected to offer similar food security, economic, environmental, and social outcomes as certified OA (UNEP-UNCTAD, 2008).

To address the above-mentioned research gaps, we applied the SMART-Farm Tool (Schader et al., 2016) to two certified organic initiatives (one in Ghana and one in Kenya) and three non-certified organic initiatives (one in Ghana and two in Kenya) that promoted OA among smallholders. By selecting a suitable control group in each study site and using entropy weights to account for the risk of selection bias, this study analyzed the effects of the five OA initiatives on the sustainability performance of the targeted group of smallholders. Importantly, we did not examine the effects associated with the uptake of OA, but the sustainability effects of the initiatives on the farmers targeted by the initiative, independently whether OA was adopted or not by the farmers. Effects were analyzed at two levels: the "sub-theme" level, an aggregate farm-level sustainability score computed through a range of indicators embedded into the SMART-Farm Tool; the "key-indicator" level, a set of specific practice, and outcome-based indicators used to calculate the sub-theme scores.

By doing so, we aimed to contribute to the ongoing debate on the effectiveness of sustainability initiatives, with a specific focus on both certified and non-certified OA initiatives (Akoyi and Maertens, 2018; Bonisoli et al., 2019). The inclusion of a diverse set of initiatives in this ex-post evaluation presents a unique opportunity to obtain a broader understanding of the impact pathway of such initiatives on the sustainability of smallholders and therefore deduce relevant policy implications to mainstream OA into national agricultural systems. The latter assumes even more importance when considering that the lack of accountability of such initiatives is regarded as a significant obstacle to increase funding for agroecological research and related development projects (IPES-Food Report, 2020). Finally, we discuss the suitability of the SMART-Farm Tool as a method for assessing the sustainability impacts of agricultural initiatives promoting OA.

2. Materials and methods

2.1. Case studies selection

The research was carried out within the Organic Farming Systems Africa (OFSA) project as a cross-sectional study whereby we collected data from a sample of farmers being exposed to existing OA initiatives and suitable control groups. The overall rationale behind selecting the case studies was to cover the relevant agro-ecological (i.e. humid and semi-arid), agronomic (i.e. predominantly arable and predominantly perennial systems), and commercial contexts (i.e. non-certified production for household consumption and local markets, and certified production for export markets) in which OA is implemented in SSA.

Kenya and Ghana were selected as focal countries primarily because both countries had a relatively high share of area under certified OA (6% and 2%) (Willer and Lernoud, 2019), and reliable local scientific partners we had previously collaborated with could implement the research project. In both countries, relevant organic initiatives (eight in Kenya, five in Ghana) were mapped and visited. The selection of the sites was then made according to the following criteria: a sufficient number of individual smallholder farms, complying with the farm selection criteria (see section 2.2); the willingness of the initiative operators to cooperate with the research team; and the coverage of different agro-ecological, agronomic and commercial contexts (Schader et al., 2021). Given this setting, the selected initiatives represented the diversity of environments in which OA was implemented to a reasonable extent. Nevertheless, we cannot claim the representativeness of the results for all initiatives implemented across SSA.

Table 1 shows the main characteristics of the five selected case studies and OA initiatives included in the study. Even though implemented in different agro-ecological zones, targeting different crops, and having different aims, all share the capacity development component on OA. Group trainings on organic management practices were provided to improve product quality, productivity, or food safety, leading to improved farm management, investments, better prices, and quality of life. The two certified initiatives also embedded the market component, as they involved paying premium prices for the certified products. Given the overlap of outcomes among the capacity and market components (ITC, 2016; Oya et al., 2017), we assumed that the following generic theory of change applies to all five initiatives: capacity development activities on OA serve as a catalyst for improving the sustainability performance of smallholder farmers.

2.2. Farm selection

As described in Schader et al., 2021, we first characterized the

Table 1
Main characteristics of the selected case studies and initiatives promoting OA in Ghana and Kenya.

	Ghana –Non certified *(GH-NC)*	***Kenya –Non certified 1*** *(KE-NC1)*	***Kenya –Non certified 2*** *(KE-NC2)*	***Ghana –Certified 1*** *(GH-C)*	**Kenya – Certified** *(KE-C)*
Country	*Ghana*	*Kenya*	*Kenya*	*Ghana*	*Kenya*
Region	*West Mamprusi*	*Murang'a*	*Machakos*	*Atwima-Mponua*	*Kirinyaga*
Agro-ecological zone	*Semi-arid*	*Humid*	*Semi-arid*	*Humid*	*Humid*
Implementing Actor(s)	*Local NGO*	*Local NGO*	*Local NGO*	*Consultancy and Producer Organization*	*Private Company and certification body*
Start of the initiative (year)	*2009*	*2006*	*2007*	*2009*	*2010*
Overall # of farmers targeted by the initiative	*NA*[a]	*10.000*	*500*	*1500*	*3.000*
Main aim of the initiative	*"To upscale environmentally friendly innovative approaches in food production to improving livelihoods of the rural poor farmers in northern Ghana"*	*"Sensitize the farming communities on the importance of organic agriculture practices for better nutrition and livelihood improvement"*	*"Train farmers on organic management practices and link them with possible market channels"*	*"Organize and train farmers in organic standards to access organic markets for cocoa?"*	*"Creating a profitable model for marketing macadamia nuts"*
Intervetion component	*Capacity Building*	*Capacity Building*	*Capacity Building*	*Capacity Building + Price premiums*	*Capacity Building + Price Premiums*
Main topics covered by capacity building activities rowhead					
Organic crop practices	✓	✓	✓	✓	✓
Post-harvest		✓	✓	✓	✓
Livestock management,				✓	✓
Financial management		✓	✓	✓	✓
Social issues e.g. child labor				✓	✓
Frequency of trainings	*Each farmer followed one or more group trainings*	*Group trainings*	*Group trainings*	*Group trainings*	*Full day trainings on different topics. Each farmers is obliged to follow at least two.*
Certification schemes adhered to	*None*	*None*	*None*	*Rainforest Alliance and Organic*	*Organic*
Certified crops	*None*	*None*	*None*	*Cocoa*	*Macadamia nuts*
Main crops grown in the case study sites	*Beans, groundnuts, maize, millets*	*Tea, beans, brassicas, maize, roots/tubers*	*Beans, maize, mango*	*Banana/plantains, cocoa, roots/tuber*	*Macadamia, coffee, maize*
Main crop targeted by the intervention	*Vegetables*	*Cabbage*	*Mango*	*Cocoa*	*Macadamia nuts*
Share of agricultural area under the targeted crop among intervention farmers	*Na*[a]	*50%*	*10%*	*26%*	*29%*
Provision of inputs (i.e. Seedlings, organic fertilizers and/or pesticides, others) or credit	*No*	*No*	*No*	*No*	*Plant protection and plant nutrition products*

[a] Not Available.

population of intervention farms in each case study area according to the socio-demographic and farm data provided by the organic initiatives and defined criteria for selection. For example, farms had a) to be located not more than 50 km away from each other, b) to be exposed to the intervention, which aimed at the adoption of OA, at least three years before the start of the data collection period in November 2016, and c) to meet or exceed the minimum farm size (KE-C: 5 macadamia trees, KE-NC1: maximum of 3 ha of farmland, KE-NC2: 5 mango trees; for the case studies in Ghana, the maximum farm size was 10 ha). In a second step, we stratified the intervention farms according to the village and randomly selected organic farms in each stratum. Finally, in a third step, we randomly selected similar control farms in each stratum. These farms needed to meet the same size criteria as the organic farms (Schader et al., 2021). By selecting the control farmers in the same villages, there is a risk of contamination. Nevertheless, we aimed at having control farms that were as much as possible comparable to the treatment farms, which explains our choice of not looking for control farmers in other areas.

Table 2
Sample sizes and number of farms according to management practices for both intervention and control groups in each study.

	GH-NC	KE-NC-1	KE-NC2	GH-C	KE-C
Total Intervention farms	***234***	***112***	***54***	***191***	***87***
Intervention farns/Organic management (Organic intended)	*38*	*46*	*15*	*58*	*80*
Intervention farms/Conventional management	*196*	*66*	*39*	*133*	*7*
Total control farms	***163***	***169***	***241***	***204***	***180***
Control farms/Organic Management (Organic-by-default)	*2*	*35*	*25*	*12*	*0*
Control farms/Conventional Management	*161*	*134*	*216*	*192*	*180*
Total case study sample size	***397***	***281***	***295***	***395***	***267***

Table 2 provides an overview of the sample sizes for both the intervention and control groups in each case study. Information on the management systems applied by the farms in each group is also included.

Table 3 provides an overview on some descriptive farm characteristics for both the control (i.e. farmers who were not exposed to the initiative) and intervention group (i.e. farmers exposed to the initiative promoting OA) (see Table 4).

Regarding the socio-economic characteristics of the analyzed farms, relatively few differences were observed. There is a significantly higher share of female-headed households in the GH-C site, and only in the KE-NC2 intervention group are farmers significantly older. The average household size is lower on intervention farms in both the GH-NC and the KE-C sites. In three out of five case studies (KE-NC-2, GH-C, KE-C), intervention farmers have more crops than control farmers. No significant differences were observed for the variables farm size and livestock ownership. Details on additional variables are reported in appendix F.

2.3. The SMART– Farm Tool and the SAFA guidelines

To evaluate the sustainability effects of the selected initiatives and their interventions in a holistic, credible, and transparent manner, we applied the *SMART-Farm* Tool (Schader et al., 2016). The method is among the most comprehensive sustainability assessment tools (Arulnathan et al., 2020; Pintér et al., 2012) and was already widely applied to assess the sustainability of individual farms in both high and low-income countries (Curran et al., 2020; Kamau et al., 2021; Landert et al., 2020; Schader et al., 2019; Ssebunya et al., 2019; Winter et al., 2020) The tool operationalizes the *Sustainability Assessment of Food and Agriculture Systems Guidelines (SAFA)* (Scialabba, 2002) developed by the *Food and Agriculture Organization of the United Nations*. The latter aimed at providing a universal and globally applicable framework for assessing sustainability. Four Sustainability Dimensions are defined: the "*Environmental Integrity*" one, with six themes and fourteen sub-themes; the "*Economic Resilience*" one, with four themes and fourteen sub-themes; the "*Social well-being*" one, with six themes and sixteen sub-themes; and the "*Good Governance*" one, with five themes and fourteen sub-themes. An overview of all themes and sub-themes is provided in Appendix A. In this manuscript, the term "sub-theme" is used interchangeably with "sustainability goal".

The SMART-Farm Tool uses a large number of indicators to compute the sustainability performance (measured on a scale of 0–100%) for each of the 58 SAFA sub-themes (on average, a number of 30.5 indicators per sub-theme) (Schader et al., 2016). Indicators can be split into generic and specific ones. Generic indicators are applied to all farm types regardless of context, whereas specific ones are context-dependent (e.g. production system, farm type, geographic location) (Curran et al., 2020). In addition, indicators can be classified into target-based, practice-based, or performance-based. Trained auditors rate all indicators based on the farm managers' answers; no measurements (e.g. soils tests) are undertaken on the farm.

Each indicator's relation to a sub-theme was expressed through a weight ranging from −1 to +1, with 0 indicating no relation among the indicator and the sub-theme existed. Indicator weights were developed in an international Delphi process involving over 60 experts from different scientific backgrounds (Schader et al., 2019). These standard weights were used in the present study.

The following equation outlines the computation of the Degree of Goal Achievement (DGA) at the farm level for each SAFA sub-theme:

$$DGA_{ix} = \sum_{n=1} (IM_{ni} \times IS_{nx}) \Big/ \sum_{n=1} (IM_{ni} \times ISmax_{n}) \forall i \text{ and } x \quad (1)$$

where x is the index of farms, i is the index of sub-themes, n is the index for the indicators that are relevant for farms of the farm type of interest and the geographical context addressed. The degree of goal achievement (DGA_{ix}) of a farm x with respect to a sub-theme i is then defined by the relation between the sum of impacts of all indicators ($n = 1$ to N) that are relevant for a sub-theme i (IM_{ni}) multiplied by the actual performance of a farm x with respect to an indicator n (IS_{nx}) and the sum of the impacts multiplied by the maximal performance possible on these indicators (IS_{maxn}). Thus, the SMART-Farm Tool can be conceptualized as a Multi-Criteria Analysis (MCA) for each sub-theme of the SAFA Guidelines (Schader et al., 2016). Being based on the SAFA framework – covering a comprehensive set of sustainability topics - enables us to test the multiple impact pathways typically assumed by initiatives promoting OA implementation.

2.4. Data collection

Trained enumerators collected data on the selected farms between November 2016 and April 2017 using the SMART-Farm Tool software. The information required to rate the indicators was collected through a 2.5 h long face-to-face interview with the farmer (i.e. the person undertaking day-to-day decisions on farm operations) and included a tour of the main fields, storage facilities, and livestock keeping areas. The

Table 3
Descriptive statistics for both control (C) and intervention (I) farms.

		GH-NC		KE-NC-1		KE-NC2		GH-C		KE-C	
		M_C	M_I	M_C	M_I	M_C	M_I	M_C	M_I	M_C	M_I
Farmer Characteristics	*Male headed holdings (%)*	0.96 (0.20)	0.97 (0.16)	0.77 (0.42)	0.78 (0.42)	0.77 (0.42)	0.74 (0.44)	0.80 (0.4)	0.73* (0.45)	0.89 (0.32)	0.90 (0.31)
	Average age household head (years)	42.4 (11.48)	41.8 (11.3)	47.6 (12.27)	47.2 (10.5)	52.2 (14.29)	57.9** (15.96)	52.6 (12.08)	54.6 (12.53)	55.9 (13.30)	56.9 (13.20)
	Average household size (count)	8.77 (3.58)	7.64*** (3.31)	2.83 (1.53)	2.66 (1.16)	4.37 (1.99)	4.43 (2.08)	5.02 (2.33)	5.47 (3.07)	2.70 (1.16)	1.71*** (0.63)
	Farmers having primary education or more (%)	0.16 (0.37)	0.30** (0.46)	0.49 (0.50)	0.55 (0.5)	0.52 (0.50)	0.56 (0.5)	0.19 (0.39)	0.20 (0.4)	0.27 (0.44)	0.15 (0.36)
	Farmers having off-farm income sources (%)	0.79 (0.41)	0.79 (0.40)	0.49 (0.5	0.42 (0.5)	0.38 (0.49)	0.43 (0.5)	0.67 (0.47)	0.69 (0.46)	0.29 (0.46)	0.34 (0.48)
Farm Characteristics	*Average farm size (ha)*	3.26 (2.38)	3.20 (2.29)	0.52 (0.51)	0.50 (0.42)	1.14 (0.84)	0.94 (0.77)	3.16 (2.35)	2.93 (2.57)	0.57 (0.51)	0.67 (0.49)
	Average # of crops grown on the farm (count)	5.70 (2.13)	5.41 (2.20)	8.74 (3.27)	8.54 (2.43)	7.79 (2.4)	9.43*** (3.04)	5.31 (1.9)	6.08*** (2.66)	5.94 (1.42)	6.73*** (1.48)
	Share of farmers owning livestock (%)	0.90 (0.3)	0.94 (0.25)	0.86 (0.35)	0.90 (0.3)	0.96 (0.19)	0.96 (0.19)	0.54 (0.5)	0.53 (0.5)	0.94 (0.24)	0.96 (0.21)

Notes: "M_C" indicates the mean for the control group; "M_I" indicates the mean for the intervention group. Standard deviations (Sd) is reported into brackets. Significance levels: *** = $p < 0.001$, ** = $p < 0.01$, * = $p < 0.05$ resulting from a two sample *t*-test for continuous variables and from a proportion test for binary variables.

Table 4
Overview of the selected sub-themes and respective goals as defined by the SAFA Guidelines.

SAFA Dimension	Selected Sub-theme	Sub-theme goal
Environmental Integrity	*Soil Quality*	*Soil characteristics provide the best conditions for plant growth and soil health, while chemical and biological soil contamination is prevented.*
	Water Quality	*The release of water pollutants is prevented and water quality is restored.*
	Species Diversity	*The diversity of wild species living in natural and semi-natural ecosystems, as well as the diversity of domesticated species living in agricultural, forestry and fisheries ecosystems is conserved and improved."*
Economic Resilience	*Stability of Production*	*Production (quantity and quality) is sufficiently resilient to withstand and be adapted to environmental, social and economic shocks*
	Stability of Market	*Stable business relationships are maintained with a sufficient number of buyers, income structure is diversified and alternative marketing channels are accessible*
	Value Creation	*Enterprises benefit local economies through employment and through payment of local taxes*
Social Well-Being	*Capacity development*	*Through training and education, all primary producers and personnel have opportunities to acquire the skills and knowledge necessary to undertake current and future tasks required by the enterprise, as well as the resources to provide for further training and education for themselves and members of their families*
	Employment relations	*Enterprises legally-binding transparent contracts with all employees that are accessible and cover the terms of work and employment is compliant with national laws on labor and social security*
	Workplace Safety	*The enterprise ensures that the workplace is safe, has met all appropriate regulations, and caters to the satisfaction of human needs in the provision of sanitary facilities, safe and ergonomic work environment, clean water, healthy food, and clean accommodation (if offered)."*
Good Governance	*Sustainability Management Plan*	*A sustainability plan for the enterprise is developed which provides a holistic view of sustainability and considers synergies and trade-offs between dimensions, including each of the environmental, economic, social and governance dimensions.*
	Conflict Resolution	*Conflicts between stakeholder interests and the enterprise's activities are resolved through collaborative dialogue (i.e. arbitrated, mediated, facilitated, conciliated or negotiated), based on respect, mutual understanding and equal power*

Source: SAFA Guidelines

information collected was directly entered into the SMART software during the interview. The data collection was performed in accordance with all relevant institutional and national ethical guidelines. Participants were asked for consent prior to the survey. Approval by an ethics committee was not required in accordance with Swiss law.

The following measures were adopted along the data collection process to ensure data quality.

1. Qualified enumerators (e.g. knowledgeable about agriculture and having prior data collection experience) were engaged to collect data.
2. Enumerators underwent two intensive weeks of training on the SMART method and were accompanied on the field during their first assessments by an experienced auditor.
3. Exchange sessions among enumerators were organized during the data collection period to ensure harmonized ratings.
4. During and after data collection, the research team performed substantial data quality checks to minimize data entry errors (more details on the data quality process are provided in Appendix E).

Given that indicator ratings primarily relied on verbal information provided by the farmers, factors undermining the credibility of the farmers' answers (e.g. fears and expectations) were mitigated. For this purpose, we undertook group sensitization meetings with the farmers and local leaderships before the commencement of the study. During these meetings, farmers were explained: a) that farm visits were aimed at collecting a large set of information related to the economic, social and environmental facets of their farming activities; b) that no economic, tangible benefits were to be expected to compensate for participation (only reports displaying the farmers' performance would be handed out at the end of the project); c) that the collected information would only be shared at an aggregated level outside of the research team in order to guarantee anonymity. Subsequently, at each assessment, the participation conditions were again clearly outlined by the enumerators before starting the interview. This process was especially relevant to avoid biased answers when discussing sensitive topics such as input use on certified organic farms or child labor.

2.5. Sub-themes and indicator selection for the current manuscript

For purposes of length, in this manuscript, only 11 out of the 58 sub-themes measured through the SMART-Farm Tool were selected for further analysis. The research team undertook the selection through a half-day prioritization workshop. All sub-themes defined by the Sustainability Assessment Of Food And Agriculture Systems (SAFA) guidelines were jointly rated against three self-defined criteria using a Likert scale ranging from 0 (not relevant) to 5 (very relevant). The considered criteria are explained below.

i Relevance within the african smallholder farming context

For instance, the sustainability goal embedded in the "Responsible Buyers" sub-theme is less relevant in the African context, given that smallholder farmers usually cannot influence input prices.

ii Suitability of the SMART-Farm Tool to measure the degree of goal achievement for the defined topic

For instance, a sub-theme such as "Greenhouse Gas Emissions" can be modeled quantitatively more accurately through other methodologies than SMART-Farm Tool.

iii Linkage with the generic theory of change of the analyzed initiatives

None of the initiatives targeted livestock production activities; therefore, a sub-theme such as "Animal Health" is less directly linked to the theory of change of the initiatives.

The results of this process are presented in Appendix C. Table 5 displays the sub-theme goals for the selected sub-themes as specified in the SAFA Guidelines.

Representing an "aggregated" figure summarizing the farmer's performance on a set of indicators, sub-theme scores alone do not provide an in-depth picture of the initiatives' effects on specific practices or other factors determining the sustainability scores. Hence, we also reported the effects of the initiatives on selected key indicators for each selected sub-theme. We identified key indicators by separating indicators into quartiles, using the absolute weights, and selected those

Table 5
Key Indicators overview.

ENVIRONMENTAL INTEGRITY
The average share of agricultural area on which insecticides (SQ_1), fungicides (SQ_2), herbicides (SQ_3) are applied; Share of farmers using active ingredients highly persistent in the soil (SQ_4). The average amount of mineral N fertilizers applied (kg/ha/year) (SW_5). Share of farmers applying compost (SQ_6), applying mulch (SQ_7). Share of farmers highly toxic to aquatic organisms (WQ_1), highly persistent in the water (WQ_2), managing the risk of nutrients pollution (WQ_3), managing riparian strips extensively (WQ_4), using active ingredients toxic to bees (SD_1).
ECONOMIC RESILIENCE
Share of farmers affected by yield loss due to pests or diseases (SP_1), affected by yield loss due to lack of access to water (SP_2), growing crop resistant varieties (SP_3), that experienced staff shortages (SP_4). The average number of buyers per farm (SM_1), the share of farm income generated by the main buyer (SM_2), the share of farmer having access to alternative markets (SM_3), having a form of cooperation with buyers (SM_4). Average weekly working hours (VC_1), share of farms on which new jobs were created (VC_2), proportion of the externally sourced inputs that is locally produced (VC_3).
SOCIAL WELL-BEING
Share of farmers that received more than one training day/year (CD_1), that have adequate access to advisory services (CD_2). Share of farmers providing written contracts to workers (ER_1), providing social protection to workers (ER_2), paying at least the regional minimum wages (ER_3), allowing workers to have regular breaks (ER_4). Share of farms on which the risk of a child performing hazardous work was considered high (ER_5), on which the risk of child labor affecting school performance was considered high (ER_6). Share of farmers using pesticides classified by the WHO as acute toxic to the health (WS_1), acutely toxic when inhaled (WS_2), having adverse long-term effects on the users (WS_3), using or providing adequate protective gear (WS_4), ensuring adequate facilities to workers (WS_5).
GOOD GOVERNANCE
Share of farmers that recognize sustainability principles (MP_1), having at least one sustainability measure planned (MP_2), aware and informed about future market challenges (MP_3), future policy changes (MP_4), climate change challenges (MP_5). Share of farmers that cooperated with other farmers (CR_1), that had conflicts over access to water (CR_2), that had conflicts over water quality (CR_3).

indicators within the highest quartile. Among these, we excluded those considered not relevant to the African context or having data quality issues. Indicators not lying within the highest quartile but having weights higher than 80% or being directly connected with OA (e.g. application of compost) were also included. An overview of the key indicators follows, while a more detailed description is found in Appendix D.

2.6. Statistical analysis

In observational studies, selection bias occurs when the reason for participation in a program correlates with the outcomes of interest. For instance, organic initiatives might target a group of farmers having certain characteristics (e.g. higher education levels), which in turn influences behavior (e.g. implementation of organic management practices). Besides exclusion bias based on specific characteristics, there is a risk of self-selection, as farmers might choose themselves whether to participate in an intervention or not. This risk is exceptionally high in the context of sustainability initiatives (Gibbon et al., 2010; Oya et al., 2017) and can complicate causal interpretations. So far, the reliability of many studies has been compromised due to the lack of suitable methods to ensure causality (De Fries et al., 2017).

To address the risk of selection bias on observable farm and farmer characteristics in the context of this study, we used entropy weights (Hainmueller and Xu, 2013). This approach offers several advantages compared to other commonly applied methods, such as propensity score matching: (i) weights are directly adjusted to the known sample moments, thereby obviating the need for continual balance checking; (ii) no observations are discarded throughout the process; (iii) the weights can be passed to any standard estimator for the subsequent estimation of treatment effects (Hainmueller, 2012). The same method was applied by Meemken and Qaim (2018a) to assess the impacts of private food standards (Fair Trade and UTZ) on gender equality.

In formal terms, let the farmer's participation in the intervention be represented by the binary treatment variable $D_i \in \{1, 0\}$, which is takes values of 1 or 0, if unit *i* is participating in the intervention or part of the control group respectively. *X* is a vector of covariates. The entropy weights *w* are assigned to each control farm *i* by the following reweighting scheme that minimizes the entropy distance metric:

$$\min_{w_i} H(w) \sum_{\{i \mid D=0\}} w_i \log\left(\frac{w_i}{q_i}\right) \tag{2}$$

subject to balance and normalizing constraints:

$$\sum_{\{i \mid D=0\}} w_i c r_i(X_i) = m_r \text{ with } r \in 1, \dots, R \tag{3}$$

$$\sum_{\{i \mid D=0\}} w_i = 1 \tag{4}$$

$$w_i \geq 0 \text{ for all } i \text{ such that } D = 0 \tag{5}$$

where $q_i = 1/n_0$ is a base weight and $cr_i(X_i) = m_r$ describes a set of *R* balance constraints imposed on the covariate moments of the reweighted control group (Hainmueller and Xu, 2013).

We calculated entropy weights based on the ebalance routine available in the software STATA 14. A rich set of covariates was used to capture farmer and farm characteristics and thus control for potential sources of bias. The covariates included age of the person managing the farm operations regularly, gender of household head, age and education level of the farmer (if not the household head), household size, years of farming experience, availability, amount of family labor, availability of off-farm income, group membership, farm size, and irrigation use. When possible, we sought to achieve balance for both mean and variance (see Appendix F for the outcomes of the reweighting).

Subsequent to the computation of weights, we used a doubly robust controlled regression model to determine the effects of the OA initiatives on the DGA and the different key outcome indicators. This model included the covariates used to compute the entropy weights as regressors as well as the entropy weights as analytical weights. Within each case study *c*, we estimated the effects on the outcome variables using ordinary least squares regression:

$$Y_{i|c} = \alpha_{0\ i|c} + w\alpha_{1i|c} INT + w\alpha_{2i|c} X + u_{i|c}$$

Where *Y* is one of the outcome variables (DGA or key indicator) for each of the farms *i* and *X* is the vector of covariates capturing observable confounders for each farm *i*. INT is the farm and intervention specific treatment effect, while *u* denotes the error term. *α* are the estimated coefficients. The estimated coefficients are interpreted as the Average Treatment Effect on the Treated (ATT) (Siebrecht, 2020), as only the untreated observations are weighted.

Additionally, we tested for heteroscedasticity using the Breusch-Pagan test (Breusch and Pagan, 1979), and integrated robust standard errors (White, 1980) to obtain unbiased standard errors of coefficients. As described, entropy balancing only controls for observed confouders, but bias mights also stem from unobservable characteristics, such as a farmer's ambition. To verify the robustness to omitted variable bias we used the approach proposed by Oster (2019), where "controlled" and "uncontrolled" regressions are compared under a set of assumptions about the relationship between selection on observables and unobservables (Bryan et al., 2020). The results (see Appendix G) show that only in few instances the results differ among the two models.

3. Results

For each sustainability dimension, first, the effects of the initiatives on the modeled sub-themes scores are described. In the text, the effects are reported into brackets and indicate the increase or decrease of the sustainability score (in percentage points), determined by the initiatives. Then, for each sub-theme, the effects on selected key indicators are reported. Here, the ATT needs to be interpreted based on the unit of the variable.

3.1. Environmental Integrity

Within the "Environmental Integrity" dimension, the sub-theme scores were significantly (p-value <0.05) improved by the OA initiatives in a substantial number of cases (9 out of 15), here intended as a case study - sub-theme configuration. As Fig. 1 shows, the "Soil Quality" score was positively affected by the KE-NC1 (+2 pp), KE-NC2 (+2 pp), the GH-C (+2 pp), and KE-C (+5 pp) initiatives. The "Water Quality" score was positively affected by the KE-NC1 (+2 pp), the GH-C (+3 pp) and KE-C (+8 pp) initiatives, and the "Species Diversity" score were positively affected by the KE-NC1 (+2 pp), the KE-NC2 (+3 pp) and KE-C (+7 pp) initiatives. Even though statistically significant, the magnitudes of the effects on the scores were generally low, reaching a maximum of 8 pp in the KE-C site for the "Water.Quality" sub-theme (see Fig. 1). Remarkably, the GH-NC initiative negatively affected the Soil Quality score (−1 pp) (see Fig. 2).

3.1.1. Sub-theme: Soil Quality

Among the indicators used to proxy the risk of chemical soil or water contamination – the area sprayed with synthetic insecticides was significantly lowered on interventions farms by the KE-NC2, GH-C, and KE-C initiatives (8, 7, and 32 pp). Likewise, the area sprayed with synthetic fungicides was significantly decreased on interventions farms by the KE-NC1 and GH-C (by 1 and 15 pp) sites. In contrast, the area sprayed with synthetic herbicides was lowered considerably only by the KE-C initiative (17% pp). Considering the application of synthetic fertilizers, intervention farmers used significantly less nitrogen fertilizers in the KE-C (−23.7 kg/ha) site. The GH-NC and KE-C initiatives increased the percentage of farmers preparing and applying compost (22% and 18% pp). None of the initiatives positively affected the amount of farmers using mulch. On the contrary, in the GH-NC site, more control farmers applied such practice. Moreover, none of the initiatives positively affected the share of farmers correctly disposing of their waste. Notably, no significant adverse effects of the initiatives on any of the indicators influencing the "Soil Quality" sub-theme were found within the KE-NC1, KE-NC2, and GH-C sites. Nevertheless, even though a significant improvement at the sub-theme score level was observed for the KE-C site, the use of active ingredients highly persistent in the soil (half-life >180 days) was negatively affected i.e. more of farmers exposed to the initiative applied active ingredients highly persistent in the soil compared to the control farms (see Appendix H, Table 1).

3.1.2. Sub-theme: Water Quality

The initiatives in the KE-NC1, KE-NC2 and the KE-C sites significantly lowered the share of farmers using active ingredients highly persistent in water (by 23, 14, and 51 pp). Similarly, the percentage of farmers using active ingredients (either slightly, moderately, highly, or very highly) toxic for aquatic organisms was lowered by the KE-NC1, KE-NC2, and KE-C (by 21, 13, and 11 pp). In the other case studies, the interventions had no significant effects. Moreover, the KE-C initiative increased the share of farmers preventing the risk of polluting water bodies (by 17 pp) and the share of farmers managing riparian strips extensively (by 17 pp). Although not significant, a negative effect is observed in the GH-NC site, where fewer intervention farmers managed riparian strips extensively (see Appendix H, Table 1).

3.1.3. Sub-theme: Species Diversity

The relatively weak performance in this sub-theme across the case studies (Fig. 1) is largely influenced by the indicators discussed within the previous sub-sections, e.g. the share of agricultural areas sprayed with pesticides, use of active ingredients toxic to aquatic organisms. Additionally, the use of active ingredients (either slightly, highly, or very highly) toxic to bees represented a critical issue affecting this goal. Except within the KE-NC1 site, these were used by a substantial share of control farmers (from 87% to 98% of farmers). Even though all interventions had a positive effect, only the KE-NC2 significantly lowererd the share of intervention farmers using active ingredients toxic to bees

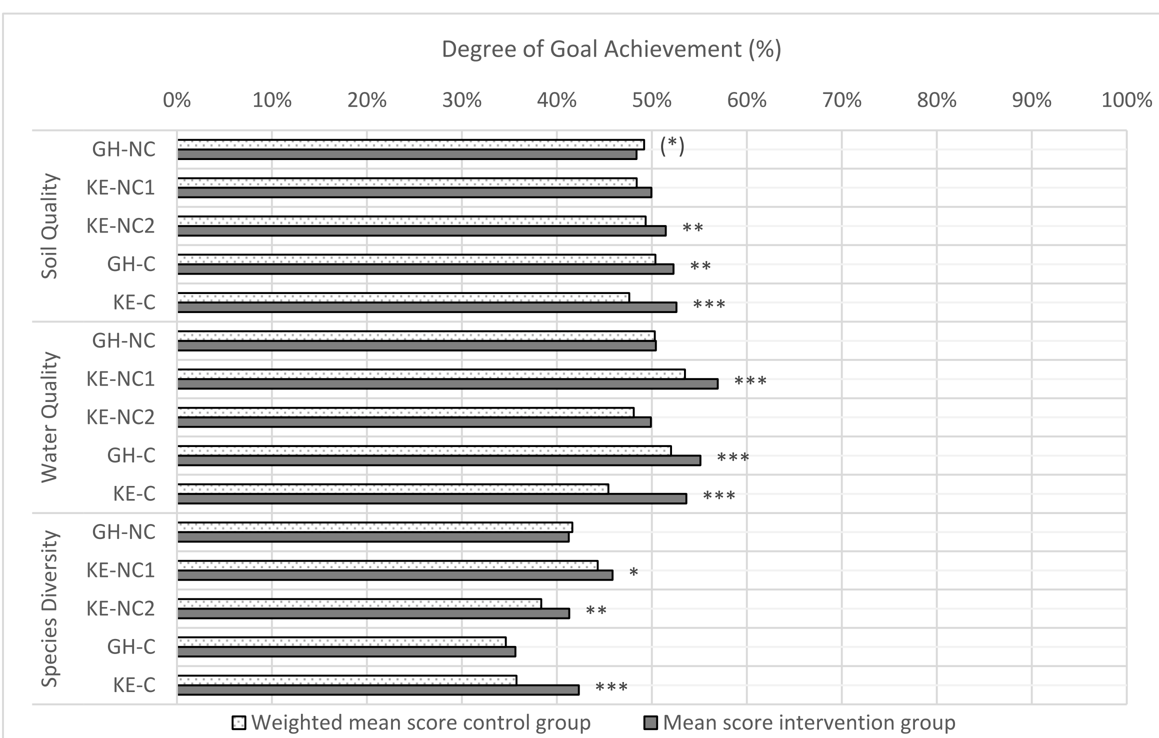


Fig. 1. Effects of the initiatives on the sustainability scores for the "Soil Quality", "Water Quality" and "Species Diversity" sub-themes in the five case studies. Significance levels: *** = p < 0.001, ** = p < 0.01, * = p < 0.05. Brackets are used to highlight negative effects.

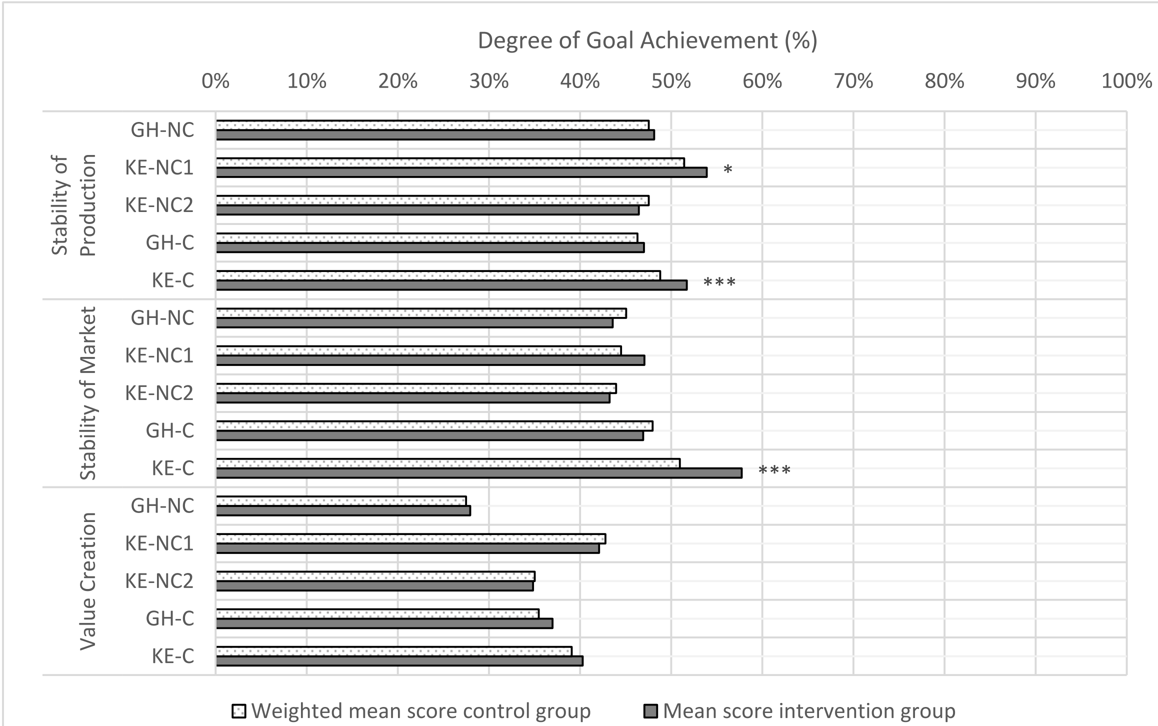


Fig. 2. Effects of the initiatives on the sustainability scores for the "Stability of Production", "Stability of Market" and "Value Creation" sub-themes in the five case studies. Significance levels: *** = $p < 0.001$, ** = $p < 0.01$, * = $p < 0.05$. Brackets are used to highlight negative effects.

(by 21 pp) (see Appendix H, Table 1).

3.2. Economic Resilience

In the "Economic Resilience" dimension, we found fewer cases where the sub-theme scores were significantly (p-value <0.05) affected by the OA initiatives (only 3 out of 15). However, the KE-C initiative positively affected both the "Stability of Market" (+7 pp) and "Stability of Production" (+3 pp) scores. The KE-NC1 also positively influenced "Stability of Production" (+2 pp) scores, while the GH-NC, KE-NC2, and GH-C initiatives did not significantly improve any sub-theme scores.

3.2.1. Sub-theme: Stability of Production

The GH-NC and KE-C initiatives decreased the share of farmers that experienced yield losses due to pests and diseases than the control farmers (by 21 and 47 pp). Nevertheless, a substantial percentage of intervention farmers still perceived high yield losses (i.e. more than 20% of the expected harvest) in these sites (respectively 23% and 18% of farmers). Additionally, the KE-C initiative lowered the share of farmers suffering from yield losses due to the lack of access to water (by 12 pp). However, in the GH-NC site, more intervention farmers were affected. Furthermore, mixed effects were observed across interventions for the variable "Growing of disease tolerant varieties" – considered a strategy to reduce yield losses and thus increase production stability. In the GH-NC site, the intervention increased the share of farmers that grew resistant varieties (by 22 pp). In contrast, in the KE-NC1, GH-C, and KE-C sites, interventions decreased the percentage of farmers growing resistant varieties (14, 12, and 24 pp). Finally, in the KE-C case study, intervention farmers experienced significantly more staff shortages than control farmers (Appendix H, Table 2).

3.2.2. Sub-theme: Stability of Market

The farms analyzed within our case studies were relatively diversified, as shown by the average number of crops grown (ranging on average between 5 and 9 crops across case studies) and the integration of livestock activities by large shares of farmers (see Table 2). However, diversification in terms of number of buyers for all farm produce is relatively limited in the GH-NC and GH-C sites, where high shares of intervention and control group farms (on average 57% and 51%) sold their produce through only one market channel. This result is explained by the fact that most producers in these sites only possess few products for sale, while the other products are generally for household consumption. Differently, in the KE-C site, farmers had on overage 4 buyers for their farm produce and a significant effect of the initiative was obseverved (i.e. initiative farmers had on average 1 market channel more for their farm produce). Moreover, in the latter case study, the initiative increased the share of farmers having access to alternative markets (33 pp). Finally, the economic dependency of farmers on the main buyer ranged from 32% of total farm income in the KE-NC2 site, to 72% in the GH-C site, and none of the initiatives had neither a positive nor a negative effect on this variable. Finally, the KE-NC1 increased the share of farmers that had a form of cooperation with their buyers (e.g. provision of inputs) (by 23 pp), while in the GH-NC site and the GH-C sites the opposite trend was observed (Appendix H, Table 2).

3.2.3. Sub-theme: Value Creation

Even though previous studies associated the conversion to organic farming with increased labor requirements (Bryan et al., 2020), none of the five initiatives were found to have increased the average weekly working hours on the farms. On the contrary, in the KE-C case study, weekly average working hours were significantly lower on intervention farms (−7.8 h per week). Furthermore, only intervention farms in the KE-C initiative significantly contributed to the creation of additional jobs at the farm level. Also, the extent to which the analyzed farms benefitted the local economies through the purchase of locally produced inputs was low (as most externally sourced inputs e.g. pesticides, fertilizers were not produced in a range of 150 km, considered as local in this study), and the interventions had neither positive nor negative effects (Appendix H, Table H2).

3.3. Social Well-Being

Within the "Social Well-Being" dimension, we found relatively few cases (5 out of 15) where the sub-theme scores were significantly (p-value <0.05) improved by the interventions (see Fig. 3). The "Workplace Safety" score was positively affected by the GH-NC (+4 pp), KE-NC1 (+3 pp), GH-C (+3 pp), and KE-C (+4 pp) interventions. The "Capacity Development" score was significantly improved by the GH-NC intervention (+4 pp), while intervention farmers in the KE-NC2 site scored significantly lower (−5 pp) compared to the control farms. Finally, the "Employment Relations" scores were significantly lower for the intervention farmers in the KE-NC2 (−3 pp) site (see Fig. 4).

3.3.1. Sub-theme: Capacity Development

The GH-NC and KE-C initiatives increased the share of intervention farmers (17 and 39 pp) that received at least one training day/year. On the other hand, in the KE-NC2 site, the intervention lowered the share of farmers that received one training day/year (22 pp less). The initiatives did not determine an increase or decrease in farmers' share with adequate access to advisory services. Such a result is surprising, especially for the certified initiatives, as field officers are supposed to facilitate the accessibility of information to the farmers. (Appendix H, Table 3).

3.3.2. Sub-theme: Employment Relations

Most key indicators affecting the "*Employment Relations*" sub-theme remained unaffected by the initiatives (Appendix H, Table 3). For instance, the lack of legally binding contracts, appropriate social protection for hired laborers, and the low salaries (compared to regional averages) remained negatively rated factors in all sites for both intervention and control farms. On the contrary, child labor was not a problematic issue within the analyzed case studies, except for the GH-NC site. In this case, the initiatives lowered the share of farms on which children's school performance was threatened by farm work (10 pp) or performed hazardous work on the farm (7 pp).

3.3.3. Sub-theme: Workplace Safety and Health Provisions

The KE-NC1, GH-C, and KE-C initiatives decreased the share of farmers using active ingredients classified as toxic when inhaled (by 20, 13, and 11 pp). Similarly, the percentage of farmers using active ingredients acutely toxic for human health was decreased by the GH-NC, KE-NC1, KE-NC2, and KE-C initiatives (by 15, 22, 13, and 11 pp). The use of chronically toxic active ingredients was also lowered by the GH-NC, KE-NC1, and KE-C (21, 16, and 51 pp). This finding assumes even more relevance considered that protective gear is not always used by farmers or workers, especially in the GH-NC, KE-NC2, GH-C, and KE-C sites. Finally, interestingly, a higher share of interventions farmers provided regular breaks to workers in the KE-NC1 site, but the opposite is observed in the KE-C site.

3.4. Good Governance

In the "Good Governance" dimension, we observed few cases of either positive or nergative effects. Intervention farmers had significantly (p-value <0.05) higher scores for the sub-theme "Sustainability Management Plan" in the GH-C (+3 pp) and KE-C (+10 pp) sites. Mixed effects were observed for the "Conflict Resolution" sub-theme. The score was positively affected by the intervention in the KE-C (+5 pp) sites, while negatively (−5 pp) by the GH-NC and KE-NC1 (−3 pp) interventions.

3.4.1. Sub-theme: Sustainability Management Plan

The KE-C initiative significantly increased the number of farmers that recognized sustainability principles (by 23 pp) and that had sustainability improvement measures planned for the future (e.g. buying water tanks, planting more trees) (by 24 pp). On the contrary, in the KE-NC2 sites, fewer intervention farmers have such measures planned. Moreover, the interventions positively affected the farmers' level of awareness on specific topics. For example, the KE-C intervention increased the share of farmers aware of market challenges (by 53 pp), the GH-C increased the share of farmers aware of climate change effects

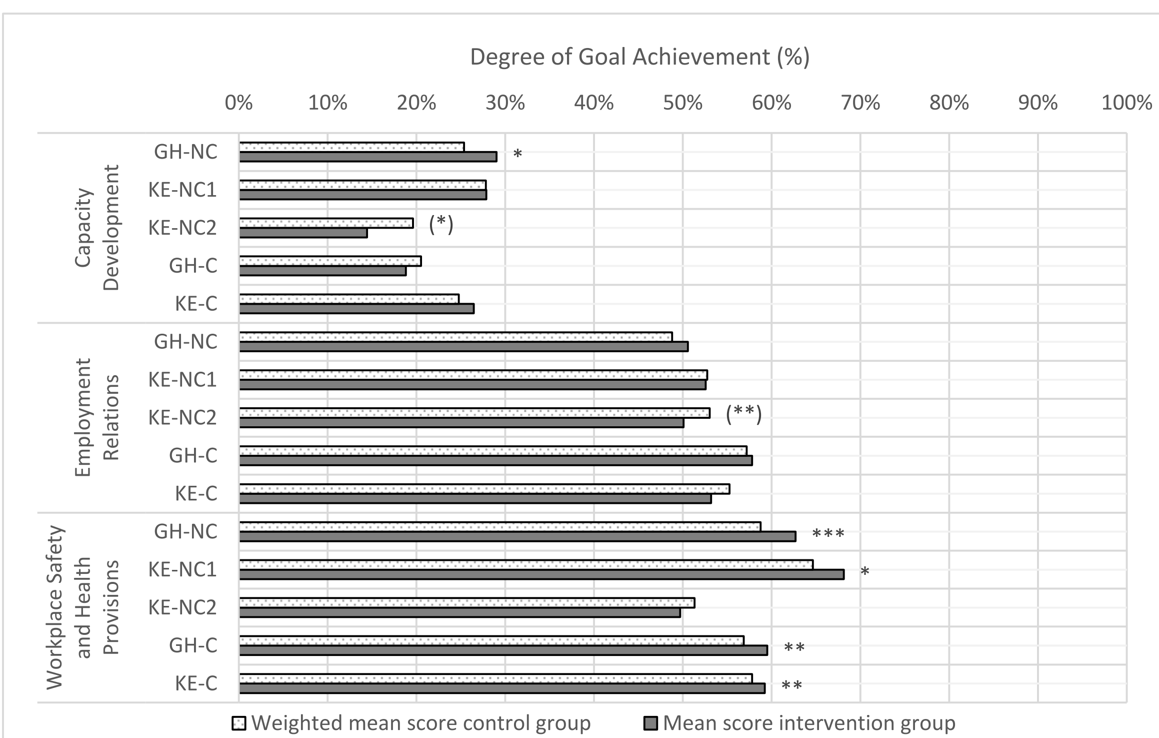


Fig. 3. Effects of the initiatives on the sustainability scores for the "Capacity Development", "Employment Relations" and "Workplace Safety" sub-themes in the case studies. Significance levels: *** = p < 0.001, ** = p < 0.01, * = p < 0.05. Brackets are used to highlight negative effects.

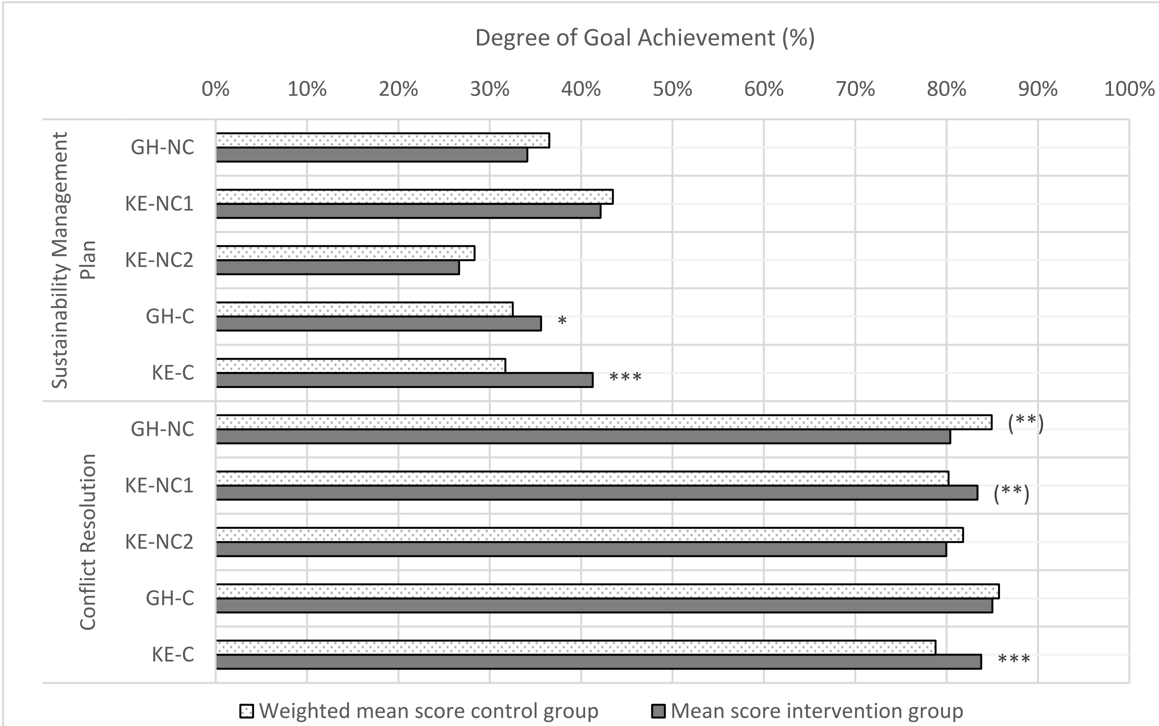


Fig. 4. Effects of the initiatives on the sustainability scores for the "Sustainability Management Plan" and "Conflict Resolution" sub-theme in all case studies. Significance levels: *** = $p < 0.001$, ** = $p < 0.01$, * = $p < 0.05$. Brackets are used to highlight negative effects.

at the farm level (by 12 pp). On the contrary, fewer intervention farmers were aware of climate change effects in the KE-NC2 site. Finally, fewer intervention farmers were aware of upcoming policy challenges in the GH-NC site (16 pp), while in the KE-C site, more intervention farmers were aware of the latter.

3.4.2. *Sub-theme: Conflict Resolution*

In the KE-NC2 site, the initiatives increased the share of farmers collaborating (i.e. exchange of tools or knowledge) with other farmers (by 16 pp). Differently, in the GH-NC site, fewer intervention farmers (7% less) had any form of cooperation with other farmers. Furthermore, conflicts over water quantity occurred relatively frequently only in the GH-NC and KE-NC2 sites, where respectively 11% and 14% of control farmers reported to have experienced conflicts over the use of water resources. In this regard, it appears the GH-NC initiative lowered the share of farmers experiencing such conflicts (8% pp), while the opposite trend is observed in the KE-NC2 site.

4. Discussion

The first part of this section discusses the findings reported in Section 3, while the second one presents lessons learnt from the application of the SMART-Farm Tool in the context of an impact evaluation of initiatives promoting OA in the African context.

4.1. *Effects of the different OA initiatives on the sustainability of smallholder farmers*

The empirical results of this farm-level assessment approach presented in Section 3 illustrated that there is no homogenous pattern in the way the initiatives influenced the sustainability performance of smallholder farmers. Whether or not significant positive or negative effects occurred varied considerably across the five initiatives and sub-theme scores. This finding is coherent with previous studies looking at the effects of certification schemes in low and middle income countries, which found that sustainability effects of organic certified or similar agroecological initiatives are highly heterogeneous based on the specific setting in which they are implemented (Oya et al., 2017). Nevertheless, we showed that the organic initiatives improved mainly the sub-theme scores within the environmental dimension, while effects within the other sustainability dimensions were more scattered. The improved performance for some key goals in the Environmental dimension, provide evidence that such initiatives have the potential to reduce negative environmental externalities. On the other hand, for those cases, where neutral or either negative effects were observed for the intervention, a deeper understanding of the impact pathways is required. Intervention effects at the indicator level followed similar trends to those at sub-theme score level, i.e. the effects tended to be initiative and outcome variable specific. Yet, the following three general trends (at indicator level) can be recognized and deserve particular attention.

4.2. *Limited influence on chemical inputs use*

Often, smallholder farmers in SSA are still associated with the term "*organic by default*" to indicate their reliance on traditional farming practices instead of externally sourced chemical inputs (Anderberg, 2020; Rosinger, 2013). However, our data shows that chemical inputs were widely used among the smallholder farmers, but to varying extents across the five study sites. Even though some of the OA initiatives reduced the agricultural area where chemical pesticides and/or mineral N fertilizers are applied, considerable shares of intervention farmers continued to rely on the latter after being trained on OA for reasons not assessed through the study. Additionally, the KE-C case study showcased that interventions might trigger unintended detrimental environmental effects due to the increased use of permitted substances under OA - such as copper-based fungicides - that are known to be persistent in the soil for more than 180 days. Moreover, the increasing use of chemical inputs by smallholder farmers in SSA (Haggblade et al., 2017; Sheahan and Barrett, 2017) indicates that conversion to OA might prove to be more demanding contrary to what previous studies suggested i.e. conversion is less risky as farmers mainly rely on traditional farming practices (UNEP-UNCTAD, 2008).

4.3. Limited influence on the adoption of organic or other agro-ecological practices

The extent to which the organic initiatives were capable of fostering the adoption of expected organic or other agro-ecological management practices e.g. application of compost, mulching, etc. was limited. Findings show few scattered positive effects of the initiatives across the case studies for the assessed practices. This result is aligned with the findings from other studies focusing on the adoption of practices by smallholder farmers in developing countries (Dietz et al., 2019; Tey et al., 2017).

4.4. Limited effects on socio-economic outcome variables

Several studies highlighted the potential of OA to positively affect various socio-economic parameters, the farms' resilience to environmental shocks (e.g. climate change), and contribute to employment generation in rural areas thanks to the higher labor requirements (e.g. due to increased labor for manual weeding). However, with the exception of positive effects on the 'Stability of production' (KE-C and KE-NCI) and 'Stability of Market' (KE-C) sub-themes, the organic interventions did not significantly influence most of the socio-economic parameters assessed in our study. These findings could be partly influenced because of the way some organic interventions were introduced into smallholder farming, i.e. targeting a certain crop, rather than the whole farm, in which case some of the effects could be limited to crop level rather than the whole farm. This assumption is reinforced by the fact that the areas under the targeted crop are relativey low, especially in the GH-NC and KE-NC1 sites. The widespread implementation gap of organic practices among interventions farms found in our study partly explains why the socio-economic outcome variables remained mostly unaffected by the interventions. Logically, changes for variables such as "Yield losses", "Full-time job equivalents per ha", "Creation of additional jobs", are strongly linked to changes in farm management practices.

4.5. Certified vs. uncertified initiatives

Given the limited sample of initiatives included in this study, a representative comparison between the sustainability effects of certified and non-certified initiatives is out of scope. Nevertheless, when undertaking a simple comparison between the effects of the two types of initiatives, we notice that certified initiatives had more frequently a sustainability effect compared to non-certified initiatives. Particularly, the KE-C initiative had by far most often a significant effect. This may be due to the fact that certification increases the likelihood that farmers will comply with the organic standards, and that certification ensure easier and more secure access to higher-value markets.

4.6. Factors explaining the limited influence of the initiatives

Several questions emerge as to the possible factors determining the limited influence of the organic interventions analyzed through the study. Even though no analysis was performed to identify and quantify these factors, we assume a mix of factors explains the continued reliance on external chemical inputs, the low uptake rates of other organic or agro-ecological management practices, and the limited sustainability effects of the interventions. Among these, the lack of awareness among intervention farmers of program rules (Dietz et al., 2019) and unclear enforcement measures might be an argument that holds true for the certified interventions, as suggested by other studies (Ansah et al., 2020; Dietz et al., 2019; Rosina Bara et al., 2018). Insufficient or inappropriate trainings on alternative pest and disease management and other agro-ecological practices (Gibbon et al., 2010) could be another reason. Moreover, initiatives often provide trainings on generic field management and post-harvest processing practices (Gibbon et al., 2010) and/or or focus solely on the certified crop (Van der Mheen-Sluijer et al., 2011). Additionally, as previous studies show, alternatives to chemical inputs (e.g. biological and/or manual control and organic fertilizers like compost) might not be readily available or used given the financial resource constraints typically faced by smallholder farmers (Bachmann, 2012), the environment they are operating in (e.g. degraded soils) (Lotter, 2015), and the lack of or limited availability of appropriate substitute inputs for organic management. Attitudes and beliefs (i.e. low expected effectiveness) or the lack of sufficient economic incentives (e.g. price premiums for organic products) might also explain the observed implementation gap. Siebrecht (2020) suggested that the top-down approach which characterizes these kinds of interventions might trigger the reactance of farmers. Another reason might be linked to the period required to see specific effects (e.g. higher yields due to improved soil management practices). On-station long-term comparative studies on organic and conventional management in Kenya have demonstrated that some of the effects of organic management occur after several years of implementation (Adamtey et al., 2016). Finally, possible conflicting initiatives from government and or private buyers (e.g. subsidizing or distributing chemical inputs) might nullify the efforts of the organic interventions.

4.7. Implications of using a holistic sustainability assessment tool for impact assessment of farm-level interventions in SSA

The coverage of a holistic set of sustainability indicators in the SMART-Farm Tool provides the potential to measure the sustainability impacts of different agricultural initiatives against a broad, well-defined, and transparent framework. The standardized set of indicators and the consequent comparability of effects among different initiatives represents a valuable feature, mainly because of the need for increased accountability of agro-ecological initiatives. However, the holistic approach entails an intrinsic trade-off regarding the level of detail - in terms of the number of indicators addressing each topic and the level of accuracy of the indicators (Ssebunya et al., 2019). This compromises the ability of the Tool to detect some relevant effects due to the lack of more targeted indicators. For instance, few indicators included in the current version of the Tool address the topic of knowledge acquisition, for which significant effects are expected given the training activities carried out by the initiatives.

Moreover, the extensive use of qualitative indicators - to allow the coverage of a wide range of topics in a cost-effective way - might compromise the ability of the Tool to capture more specific quantitative effects (e.g. on aspects such as "profitability" or reduced pesticide usage). In addition, being a farm-level assessment tool entails additional caveats. First, organic initiatives in SSA often target a specific crop or cropping system or even one single plot (Van der Mheen-Sluijer et al., 2011). Therefore, applying a farm-level assessment tool might lead to a blurring of the effects considering that assessment indicators are not crop or plot specific. Secondly, organic initiatives might, or might not, have effects beyond the farm level in the short or long term (e.g. employment creation and empowerment of women at the producer organization level), which are not captured with this approach.

5. Conclusions

The analyzed initiatives promoting OA through their interventions mainly improved smallholder farmers' sustainability performance for the environmental goals assessed. However, significant effects were less frequent regarding the other sustainability dimensions (i.e. economic, social, and good governance). Overall, although significant, the magnitude of effects was often relatively small. The inability of the initiatives to trigger change for most indicators and among the majority of targeted farmers can explain this. Notably, the "Stability of Production" (in KE-C and KE-NCI) and "Stability of Markets" (in KE-C) are worth highlighting as this may reflect the potential of OA initiatives to impact the economic resilience of the targeted farmers. Moreveor, despite the limited sample of certified and uncertified initiatives that we analyzed, the former appear to be more likely to trigger positive sustainability effects. We consider further research on the differences between these two types of initiatives is required to gain more in-depth insights on this aspect.

The overall limited influence on sustainability suggests that the impact pathway organic initiatives commonly rely on (i.e. knowledge sharing leads to adoption and associated sustainability benefits) is jeopardized by a complex set of factors. This, in turn, suggests that to successfully induce the adoption of organic practices and unlock positive environmental and socio-economic effects, initiatives must complement knowledge-sharing activities with other support measures. For example, farmers must be supported to overcome the motivational, financial, and other resource barriers explaining the OA implementation gap. Additionally, taking a farm-level approach (instead of a crop level one) could represent a way to broaden the impacts of the interventions. Nevertheless, for such impacts to occur additional measures such as the development of local organic markets would be required. In fact, local demand for organic products in developing countries is currently limited (Meemken and Qaim, 2018b). Moreover, the future inclusion of case study sites with varying lengths of organic agriculture implementation could help determine some of the long-term impacts of such initiatives on-farm sustainability, given the long learning curve for organic management. Finally, this study shows that the SMART-Farm Tool holds the potential to undertake holistic impact evaluations of initiatives promoting OA or other agro-ecological practices. It enables the assessment of initiatives against a transparent framework and offers comparability between evaluations. Therefore, it could play a relevant role in increasing the accountability of sustainability initiatives, a key point to ensure increased funding of agroecology (IPES-Food Report, 2020). At the same time, it needs to be complemented with additional indicators specific to SSA, or approaches must be integrated in the future to capture effects occurring beyond the farm level.

Credit author statement

We the undersigned declare that this manuscript is original, has not been published before and is not currently being considered for publication elsewhere.We confirm that the manuscript has been read and approved by all named authors and that there are no other persons who satisfied the criteria for authorship but are not listed.We further confirm that the order of authors listed in the manuscript has been approved by all of us.We understand that the Corresponding Author is the sole contact for the Editorial process. He/she is responsible for communicating with the other authors about progress, submissions of revisions and final approval of proofs.Signed by all authors as follows:Johan Blockeel, Christian Schader, Anja Heidenreich, Christian Grovermann, Irene Kadzere, Irene S. Egyir, Anne Muriuki, Joseph Bandanaa, Chrysantus M. Tanga, Joseph Clottey, John Ndungu, Matthias Stolze.

Funding sources

This work was undertaken within the Organic Farming Systems Africa (OFSA) and the ProEcoAfrica projects financed by the Mercator Foundation Switzerland, the Swiss Agency for Development and Cooperation (SDC) and the Dutch Humanist Institute for Development and Cooperation (Hivos). The donors were not involved in the writing-up of this manuscript.

Data availability

Data will be made available on request.

Acknowledgements

The authors thank all local partners in Ghana and Kenya and the enumerators and other field staff for contributing to the success of this study. Finally, we thank all farmers for their invaluable participation in this study and for sharing relevant data with us. We also acknowledge the guidance from the project advisory members in implementing the project.

Appendices

A. SAFA Framework: Overview of themes and sub-themes for each sustainability dimension

GOOD GOVERNANCE					
CORPORATE ETHICS	Mission Statement	Due Diligence			
ACCOUNTABILITY	Holistic Audits	Responsibility	Transparency		
PARTICIPATION	Stakeholder Dialogue	Grievance Procedures	Conflict Resolution		
RULE OF LAW	Legitimacy	Remedy, Restoration & Prevention	Civic Responsibility	Resource Appropriation	
HOLISTIC MANAGEMENT	Sustainability Management Plan	Full-Cost Accounting			

ENVIRONMENTAL INTEGRITY			
ATMOSPHERE	Greenhouse Gases	Air Quality	
WATER	Water Withdrawal	Water Quality	
LAND	Soil Quality	Land Degradation	
BIODIVERSITY	Ecosystem Diversity	Species Diversity	Genetic Diversity
MATERIALS & ENERGY	Material Use	Energy Use	Waste Reduction & Disposal
ANIMAL WELFARE	Animal Health	Freedom from Stress	

ECONOMIC RESILIENCE					
INVESTMENT	Internal Investment	Community Investment	Long-Ranging Investment	Profitability	
VULNERABILITY	Stability of Production	Stability of Supply	Stability of Market	Liquidity	Risk Management
PRODUCT QUALITY & INFORMATION	Food Safety	Food Quality	Product Information		
LOCAL ECONOMY	Value Creation	Local Procurement			

SOCIAL WELL-BEING				
DECENT LIVELIHOOD	Quality of Life	Capacity Development	Fair Access to Means of Production	
FAIR TRADING PRACTICES	Responsible Buyers	Rights of Suppliers		
LABOUR RIGHTS	Employment Relations	Forced Labour	Child Labour	Freedom of Association & Right to Bargaining
EQUITY	Non Discrimination	Gender Equality	Support to Vulnerable People	
HUMAN SAFETY & HEALTH	Workplace Safety and Health Provisions	Public Health		
CULTURAL DIVERSITY	Indigenous Knowledge	Food Sovereignty		

.

B. Sustainability objectives as defined by SAFA for the 11 selected sub-themes

SAFA Dimension	Selected Sub-theme	Sub-theme goal
Environmental Integrity	*Soil Quality*	*Soil characteristics provide the best conditions for plant growth and soil health, while chemical and biological soil contamination is prevented.*
	Water Quality	*The release of water pollutants is prevented and water quality is restored.*
	Species Diversity	*The diversity of wild species living in natural and semi-natural ecosystems, as well as the diversity of domesticated species living in agricultural, forestry and fisheries ecosystems is conserved and improved."*
Economic Resilience	*Stability of Production*	*Production (quantity and quality) is sufficiently resilient to withstand and be adapted to environmental, social and economic shocks*
	Stability of Market	*Stable business relationships are maintained with a sufficient number of buyers, income structure is diversified and alternative marketing channels are accessible*
	Value Creation	*Enterprises benefit local economies through employment and through payment of local taxes*
Social Well-Being	*Capacity development*	*Through training and education, all primary producers and personnel have opportunities to acquire the skills and knowledge necessary to undertake current and future tasks required by the enterprise, as well as the resources to provide for further training and education for themselves and members of their families*
	Employment relations	*Enterprises legally-binding transparent contracts with all employees that are accessible and cover the terms of work and employment is compliant with national laws on labor and social security*
	Workplace Safety	*The enterprise ensures that the workplace is safe, has met all appropriate regulations, and caters to the satisfaction of human needs in the provision of sanitary facilities, safe and ergonomic work environment, clean water, healthy food, and clean accommodation (if offered)."*
Good Governance	*Sustainability Management Plan*	*A sustainability plan for the enterprise is developed which provides a holistic view of sustainability and considers synergies and trade-offs between dimensions, including each of the environmental, economic, social and governance dimensions.*

(continued on next page)

(*continued*)

SAFA Dimension	Selected Sub-theme	Sub-theme goal
	Conflict Resolution	*Conflicts between stakeholder interests and the enterprise's activities are resolved through collaborative dialogue (i.e. arbitrated, mediated, facilitated, conciliated or negotiated), based on respect, mutual understanding and equal power*

Source: SAFA Guidelines 2013

C. Author's evaluation of sub-themes based on criteria's presented in section 2.1.5

DIMENSION	SAFA SUB-THEME	RELEVANCE IN THE SMALLHOLDER CONTEXT (1=very low 2=low 3= good 4=high 5=very high)	SUITABILITY OF SMART APPROACH (1=very low 2=low 3= good 4=high 5=very high)	EXPECTED EFFECT OF THE INTERVENTION ON THE SUB-THEME (1=very low 2=low 3= good 4=high 5=very high)	Sum of Points for all criteria
ENVIRONMENTAL INTEGRITY	Greenhouse Gases	2	1	2	5
	Air Quality	2	1	2	5
	Water Withdrawal	3	1	2	6
	***Water Quality**	4	3	5	12
	***Soil Quality**	5	3	5	13
	Land Degradation	5	2	4	11
	Ecosystem Diversity	4	2	4	10
	***Species Diversity**	4	3	5	12
	Genetic Diversity	4	3	4	11
	Material Use	2	2	2	6
	Energy Use	2	2	2	6
	Waste Reduction & Disposal	3	3	2	8
	Animal Health	4	4	3	11
	Freedom from Stress	4	4	3	11
ECONOMIC RESILIENCE	Internal Investment	2	2	2	6
	Community Investment	3	3	3	9
	Long-Ranging Investment	3	3	4	10
	Profitability	5	1	4	10
	***Stability of Production**	4	3	4	11
	Stability of Supply	4	2	3	9
	***Stability of Market**	4	3	4	11
	Liquidity	5	2	3	10
	Risk Management	4	2	4	10
	Food Safety	5	1	4	10
	Food Quality	5	1	4	10
	Product Information	3	2	3	8
	***Value Creation**	4	3	4	11
	Local Procurement	3	3	2	8
SOCIAL WELL-BEING	Quality of Life	4	3	3	10
	***Capacity Development**	5	3	5	13
	Fair Access to Means of Production	4	2	4	10

	Responsible Buyers	3	2	1	6
	Rights of Suppliers	3	2	1	6
	***Employment Relations**	4	3	4	11
	Forced Labour	2	2	1	5
	Child Labour	4	2	3	9
	Freedom of Association and Right to Bargaining	3	2	1	6
	Non Discrimination	3	2	3	8
	Gender Equality	5	2	3	10
	Support to Vulnerable People	3	2	1	6
	***Workplace Safety and Health Provisions"**	4	4	5	13
	Public Health	4	3	3	10
	Indigenous Knowledge	2	3	3	8
	Food Sovereignty	2	2	3	7
GOOD GOVERNANCE	**Mission Statement**	3	3	2	8
	Due Diligence	3	2	2	7
	Holistic Audits	2	3	2	7
	Responsibility	2	3	2	7
	Transparency	3	2	2	7
	Stakeholder Dialogue	3	2	3	8
	Grievance Procedures	2	3	1	6
	***Conflict Resolution**	3	3	3	9
	Legitimacy	3	2	3	8
	Remedy, Restoration & Prevention	2	3	2	7
	Civic Responsibility	1	3	2	6
	Resource Appropriation	1	3	1	5
	***Sustainability Management Plan**	4	3	4	11
	Full-Cost Accounting	1	3	1	5

Note: the sub-themes selected for a more in-depth analysis in this study were marked with an asterisk "*".

.

D. Key Indicators Description

Appendix D. Table 1
Environmental Integrity – Description of key indicators

Indicator ID	Indicator Name	Question assessed by the Auditor	Assessment Scale Description
SQ_1	Average share of agricultural area on which insecticides are applied	What proportion of the agricultural area receives synthetic chemical insecticide applications?	% of agricultural area
SQ_2	Average share of agricultural area on which fungicides are applied	What proportion of the agricultural area receives synthetic chemical fungicide applications?	% of agricultural area
SQ_3	Average share of agricultural area on which herbicides are applied	What proportion of the agricultural area receives synthetic chemical herbicide applications?	% of agricultural area
SQ_4	Share of farmers using active ingredients highly persistent in the soil	Are active substances used, which are considered to be very persistent in soil (half-life >180 days) according to the "PAN Pesticide Database"?	0% = At least one active ingredient very persistent in soil is used 100% = No active ingredients very persistent in soil are used
SQ_5	Average amount of Mineral N fertiliser applied	What is the farm's mineral N fertiliser (in kg) usage per hectare per year on its agricultural area?	kg/Ha
SQ_6	Share of farmers preparing and applyng compost	Is compost applied on the fields?	0% = No 100% = Yes
SQ_7	Share of farmers disposing waste correctly	Is all operational/commercial waste disposed correctly?	0% = No 100% = Yes

Appendix D. Table 2
Economic resilience – Description of key indicators

Indicator ID	Indicator Name	Question assessed by the Auditor	Assessment Scale Description
SP_1	Share of farmers affected by yield loss due to pests or diseases	Has the farm been affected by crop failures (>20% of expected yields) in the past 5 years?	0% = Yes 100% = No
SP_2	Share of farmers affected by yield loss due to lack of access to water	Were yields limited during the past 5 years as a result of lack of water (more than 20% reduction either in livestock and/or crop production yield)?	100% = No 0% = Yes
SP_3	Share of farmers growing crop resistant varieties	Are cultivars chosen with a view to resistance to harmful organisms and diseases?	0% = No 100% = Yes
SP_4	Share of farmers that experienced staff shortages	Have there been staff shortages in the last five years that could not be resolved?	0% = Yes 100% = No
SM_1	Average number of buyer per farm	To how many buyers does the farm sell all of its products ?	# of buyers
SM_2	Share of income generate by the main buyer	What proportion of sales profit, in terms of income, does the most important buyer generate?	Share of farm income
SM_3	Share of farmer having access to alternative markets	Are there alternative markets for all products if buyers drop out?	0% = No 00% = Yes
SM_4	Share of farmers having a form of cooperation with buyers	Is there any cooperation between the farm and it's buyers beyond the commercial relationship of pure selling/buying of the farm's produce? (Examples: Common product development, production planning, advisory service from the buyer, etc.)	0% = No 100% = Yes
VC_1	Full-time jobs per hectare	The farms total full-time job equivalents are computed based on the total working days of the farmer and workers.	FTE/ha agricultural area
VC_2	Share of farms on which new jobs were created	Were there any new (= additional) jobs created or job cuts at the farm in the past 5 years?	0% = No 100% = Yes
VC_3	Proportion of the externally sourced inputs that is locally produced	What proportion of the externally sourced inputs is locally produced? (Consider the five most important inputs (in terms of cost). Rate negative if the origin of production is not known.	% of externally sourced inputs

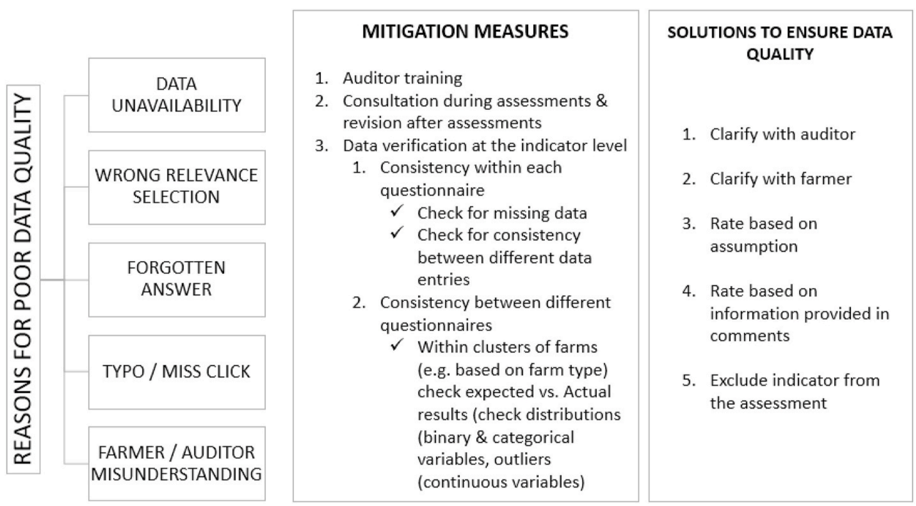
Appendix D. Table 3
Economic resilience – Description of key indicators

Indicator ID	Indicator Name	Question assessed by the Auditor	Assessment Scale Description
CD_1	Share of farmers that received more than 1 day of training/year	How many days of further education or training (per person) were taken during the last year?	Total number of days spent for further training during last year on average per person.
CD_2	Share of farmers having adequate access to extension services	Does the farm have adequate access to extension services?	0% = No 100% = Yes
ER_1	Share of farmers providing written contracts to workers	Do all of the permanent workers have legally binding employment contracts?	0% = No 100% = Yes
ER_2	Share of farmers having or providing social protection to workers	Do all workers have social protection (including injury, illness and maternity benefits) and are protected against dismissal?	0% = No 100% = Yes
ER_3	Share of farmers paying more than the regional minimum wages	How does the minimum wage that the operation pays to its workers compare with the statutory minimum wage that is generally paid in this region and sector?	0% = Lower 100% = Higher
ER_4	Share of farmers allowing regular breaks	Are all workers free to take regular breaks? - Long break of 1 h per working day (9h) - Additional breaks (at least 0.25 h) for each half day	0% = No 100% = Yes
ER_5	Share of farms on which the risk of child performing hazarduous work was considered high	(If children (<16 years old) help with the work) Is there a risk that this work may be hazardous to their health or development (e.g. carrying heavy loads or doing dangerous work, applying plant protection products)? If no children help, do not assess.	0% = Yes 100% = No
ER_6	Share of farms on which the risk of child labor affecting school performance was considered high	(If children (<16 years old) help with the work) Is there a risk that the children's school performance is hampered by that work (e.g. they are tired at school or do not have time to complete homework assignments)? If no children help, do not assess.	0% = Yes 100% = No
WS_1	Share of farmers using pesticides classified as acute toxic to the health	Are active substances used, which are classified by the WHO as acute toxic to the health of the users?	0% = Yes, at least one 100% = None
WS_2	Share of farmers using pesticides classified as acutely toxic when inhaled	Are active substances used, which are considered acute toxic when inhaled by the users according to the "Globally Harmonized System of Classification (GHS)"?	0% = Yes, at least one 100% = None
WS_3	Share of farmers using pesticides having adverse long term effects	Are active substances used, which are considered to have adverse long term effects on the users according to the "PAN List of HHPs" or "PAN Pesticide Database"?	00% = Yes, at least one 100% = None
WS_4	Share of farmers using or providing adequate protective gear	Does the farmer ensure that workers have appropriate protection during their application of pesticides and other hazardous materials?	0% = No 100% = Yes
WS_5	Share of farmers ensuring adequate facilities to workers	Are all workers able to have regular meals, drink sufficiently and use toilet facilities?	0% = No 100% = Yes

Appendix D. Table 3
Good Governance – Description of key indicators

Indicator ID	Indicator Name	Question assessed by the Auditor	Assessment Scale Description
MP_1	Share of farmers that recognized sustainability principles	Is the farm manager committed to the principles of sustainability?	0% = No 100% = Yes
MP_2	Share of farmers having sustainability measures planned	Is the farm manager able to describe planned sustainability improvements in detail?	0% = No 100% = Yes
MP_3	Share of farmers aware and informed about future market challenges	Is the farmer aware and informed about future market challenges?	0% = No 100% = Yes
MP_4	Share of farmers aware and informed about future policy changes	Is the farmer aware and informed about future policy changes/political challenges?	0% = No 100% = Yes
MP_5	Share of farmers aware and informed about climate change challenges	Does the farm manager know the predictions for climate changes in the region resulting from global climate change and can he judge the impact on the farm?	0% = No 100% = Yes
CF_1	Share of farmers that cooperated with other farmers	Does the farm successfully, and in the long term, cooperate with other farms?	0% = No 100% = Yes
CF_2	Share of farmers that had conflicts over access to water	Are there, or have there been, any conflicts with other water users/stakeholders in the farm's neighbourhood over access to water and/or the volume of water used?	100% = No 0% = Yes
CF_3	Share of farmers that had conflicts over water quality	Are there, or have there been, any conflicts with other water users/stakeholders over water quality in the farm's neighbourhood?	100% = No 0% = Yes

E. Data quality process



.

F. Entropy Balancing Weights

Appendix F. Table 1
Outcomes of the entropy balancing reweighting: GH-NC site

Covariates	Treatment Group		Control Group: Before weighting		Control Group: After weighting	
	Mean	Variance	Mean	Variance	Mean	Variance
Farm size	3.222	5.251	3.275	5.668	3.222	5.252
Irrigation use	0.172	0.143	0.244	0.186	0.173	0.144
Rented land	0.176	0.119	0.276	0.156	0.176	0.119
Household size	7.671	10.990	8.768	12.760	7.673	11.000
Family labor availability	73.49	5′997.00	66.01	5′916.00	73.47	5′990.00
Family labor proportion	0.820	0.027	0.860	0.020	0.820	0.027
Proportion of female household members	0.492	0.022	0.486	0.021	0.492	0.022
Age of farm manager	41.98	124.70	42.36	128.70	41.98	124.60
Gender of household head	1.974	0.025	1.957	0.041	1.974	0.025
Gender of farm manager	0.207	0.165	0.287	0.206	0.207	0.165
Formal education of farm manager	0.349	0.228	0.238	0.182	0.349	0.229
Farming experience of farm manager	17.59	107.40	18.77	109.70	17.59	107.30
Availability of off-farm income	1.802	0.160	1.780	0.172	1.802	0.160
Farmer organization membership	0.728	0.199	0.628	0.235	0.728	0.199

Appendix F. Table 2
Outcomes of the entropy balancing reweighting: KE-NC1

Covariates	Treatment Group		Control Group: Before weighting		Control Group: After weighting	
	Mean	Variance	Mean	Variance	Mean	Variance
Farm size	0.509	0.1861	0.5225	0.2602	0.509	0.1861
Irrigation use	0.3628	0.2332	0.3452	0.2274	0.3628	0.2326
Rented land	0.01738	0.006472	0.03166	0.01329	0.0174	0.00647
Household size	2.65	1.334	2.839	2.342	2.65	1.334
Family labor availability	1173	830,521	1215	2,161,643	1173	830,940
Family labor proportion	0.6028	0.1038	0.5961	0.1199	0.6028	0.1038
Proportion of female household members	0.5568	0.06176	0.5308	0.06111	0.5568	0.06176
Age of farm manager	47.42	113.9	47.55	147.4	47.42	113.9
Gender of household head	1.779	0.1738	1.774	0.1761	1.779	0.1733
Gender of farm manager	0.7522	0.1881	0.5536	0.2486	0.7522	0.1875
Formal education of farm manager	0.9912	0.00885	0.9762	0.02338	0.9912	0.00882
Farming experience of farm manager	22.56	163	21.74	140.1	22.56	163
Availability of off-farm income	1.416	0.2451	1.488	0.2514	1.416	0.2444
Farmer organization membership	0.823	0.147	0.7202	0.2027	0.823	0.1465

Appendix F. Table 3
Outcomes of the entropy balancing reweighting: KE-NC2

Covariates	Treatment Group		Control Group: Before weighting		Control Group: After weighting	
	Mean	Variance	Mean	Variance	Mean	Variance
Farm size	0.9425	0.6	1.138	0.7123	0.9425	0.6
Irrigation use	0.2407	0.1862	0.2552	0.1909	0.2408	0.1836
Rented land	0.0004458	0.0000107	0.004862	0.001933	0.0004526	0.000074
Household size	4.431	4.329	4.367	3.947	4.431	4.329
Family labor availability	124.3	9219	103	9214	124.3	9219
Family labor proportion	0.6944	0.07351	0.6123	0.07625	0.6944	0.07351
Proportion of female household members	0.5271	0.03978	0.4978	0.04559	0.5271	0.03978
Age of farm manager	57.87	254.7	52.24	202.6	57.87	254.7
Gender of household head	1.741	0.1957	1.766	0.1802	1.741	0.1928
Gender of farm manager	0.6296	0.2376	0.5063	0.251	0.6296	0.2342
Formal education of farm manager	0.8333	0.1415	0.9414	0.05538	0.8333	0.1395
Farming experience of farm manager	31.13	220.2	26.05	199.2	31.13	220.2
Availability of off-farm income	1.426	0.2491	1.381	0.2368	1.426	0.2455
Farmer organization membership	0.7407	0.1957	0.6025	0.2405	0.7407	0.1929

Appendix F. Table 4
Outcomes of the entropy balancing reweighting: GH-C

Covariates	Treatment Group		Control Group: Before weighting		Control Group: After weighting	
	Mean	Variance	Mean	Variance	Mean	Variance
Farm size (ha)	2.927	6.565	3.155	5.501	2.927	6.565
Rented land	0.069	0.037	0.102	0.072	0.069	0.037
Household size	5.484	9.359	5.017	5.419	5.484	9.359
Family labor availability	74.95	4′090.00	68.66	3′247.00	74.94	4′090.00
Family labor proportion	0.628	0.032	0.644	0.029	0.628	0.032
Proportion of female household members	0.491	0.055	0.491	0.040	0.491	0.055
Age of farm manager	54.54	156.30	52.70	145.50	54.54	156.30
Gender of household head	1.724	0.201	1.798	0.162	1.724	0.201
Gender of farm manager	0.469	0.250	0.419	0.245	0.469	0.250
Formal education of farm manager	0.781	0.172	0.724	0.201	0.781	0.172
Farming experience of farm manager	26.20	151.10	24.43	111.50	26.20	151.10
Availability of off-farm income	1.693	0.214	1.680	0.219	1.693	0.214
Farmer organization membership	0.703	0.210	0.409	0.243	0.703	0.210

Appendix F. Table 5
Outcomes of the entropy balancing reweighting: KE-C.

Covariates	Treatment Group		Control Group: Before weighting		Control Group: After weighting	
	Mean	Variance	Mean	Variance	Mean	Variance
Farm size	0.679	0.243	0.571	0.260	0.679	0.243
Irrigation use	0.118	0.105	0.337	0.225	0.116	0.103

(*continued on next page*)

Appendix F. Table 5 (*continued*)

Covariates	Treatment Group		Control Group: Before weighting		Control Group: After weighting	
	Mean	Variance	Mean	Variance	Mean	Variance
Rented land	0.003	0.001	0.015	0.008	0.003	0.002
Family labor availability	416.00	133′363.00	330.90	71′203.00	416.20	133′360.00
Family labor proportion	0.636	0.091	0.568	0.068	0.637	0.091
Proportion of female household members	0.499	0.109	0.520	0.040	0.499	0.108
Age of farm manager	57.32	175.90	56.04	184.10	57.32	176.20
Gender of household head	1.894	0.096	1.895	0.094	1.894	0.095
Gender of farm manager	0.282	0.205	0.204	0.164	0.283	0.204
Formal education of farm manager	0.929	0.066	0.823	0.146	0.930	0.066
Farming experience of farm manager	31.89	159.00	33.03	214.20	31.89	159.20
Availability of off-farm income	1.365	0.235	1.287	0.206	1.364	0.233
Farmer organization membership	1.000	0.000	0.917	0.076	1.000	0.000

G. Effects of the initiatives at the Sub-theme level

	Uncontrolled			Controlled					
	Effect	SE	R2	Effect	SE	R2	Beta	Bounds on the treatment effect (δ = 1, Rmax = 1.3*R)	Treatment effect excludes 0
Soil Quality									
GH-NC	−0.008	0.004	0.013	−0.009*	0.004	0.108	−0.0092	(-0.0092,-0.0088)	YES
KE-NC1	0.016	0.009	0.017	0.014	0.009	0.169	0.0132	(0.0132,0.0138)	YES
KE-NC2	0.021**	0.007	0.052	0.021**	0.007	0.217	0.0206	(0.0206,0.0208)	YES
GH-C	0.019**	0.006	0.029	0.019**	0.006	0.059	0.0193	(0.0192,0.0193)	YES
KE-C	0.05***	0.010	0.144	0.054***	0.010	0.233	0.0578	(0.0535,0.0578)	YES
Water Quality									
GH-NC	0.001	0.004	0.000	0.001	0.004	0.107	0.0003	(0.0003,0.0005)	YES
KE-NC1	0.035**	0.010	0.052	0.035***	0.009	0.171	0.0351	(0.0349,0.0351)	YES
KE-NC2	0.018	0.011	0.018	0.017	0.010	0.157	0.0168	(0.0168,0.0171)	YES
GH-C	0.031***	0.008	0.047	0.031***	0.008	0.088	0.0307	(0.0307,0.0308)	YES
KE-C	0.082***	0.012	0.260	0.091***	0.011	0.336	0.1082	(0.0911,0.1082)	YES
Species Diversity									
GH-NC	−0.004	0.005	0.002	−0.004	0.005	0.131	−0.0034	(-0.0035,-0.0034)	YES
KE-NC1	0.016*	0.007	0.020	0.015*	0.007	0.126	0.0150	(0.015,0.0152)	YES
KE-NC2	0.029**	0.010	0.061	0.029**	0.009	0.237	0.0294	(0.0294,0.0294)	YES
GH-C	0.01	0.006	0.010	0.01	0.005	0.066	0.0101	(0.0101,0.0101)	YES
KE-C	0.066***	0.011	0.193	0.073***	0.011	0.266	0.0833	(0.0725,0.0833)	YES
Stability of production									
GH-NC	0.006	0.006	0.003	0.006	0.006	0.084	0.0055	(0.0055,0.0056)	YES
KE-NC1	0.025**	0.009	0.041	0.021*	0.009	0.185	0.0191	(0.0191,0.0207)	YES
KE-NC2	−0.011	0.010	0.008	−0.011	0.009	0.125	−0.0109	(-0.0109,-0.0109)	YES
GH-C	0.007	0.007	0.004	0.007	0.006	0.128	0.0070	(0.007,0.007)	YES
KE-C	0.029**	0.010	0.059	0.041***	0.010	0.215	0.0471	(0.0411,0.0471)	YES
Stability of Market									
GH-NC	−0.015*	0.007	0.014	−0.014	0.007	0.104	−0.0131	(-0.0135,-0.0131)	YES
KE-NC1	0.026	0.014	0.020	0.026	0.014	0.195	0.0257	(0.0256,0.0257)	YES
KE-NC2	−0.007	0.013	0.002	−0.007	0.011	0.219	−0.0070	(-0.007,-0.007)	YES
GH-C	−0.01	0.011	0.004	−0.011	0.010	0.103	−0.0105	(-0.0105,-0.0105)	YES
KE-C	0.068***	0.014	0.124	0.078***	0.015	0.244	0.0865	(0.0782,0.0865)	YES
Value Creation									
GH-NC	0.005	0.008	0.001	0.005	0.008	0.087	0.0051	(0.005,0.0051)	YES
KE-NC1	−0.007	0.008	0.003	−0.007	0.008	0.106	−0.0073	(-0.0073,-0.0072)	YES
KE-NC2	−0.002	0.009	0.000	−0.002	0.009	0.069	−0.0018	(-0.0018,-0.0018)	YES
GH-C	0.015	0.008	0.012	0.015	0.008	0.042	0.0149	(0.0149,0.0149)	YES
KE-C	0.012	0.010	0.009	0.011	0.012	0.085	0.0100	(0.01,0.0106)	YES
Capacity Development									
GH-NC	0.037*	0.018	0.013	0.037*	0.018	0.077	0.0371	(0.0371,0.037)	YES
KE-NC1	0.001	0.023	0.000	0.001	0.023	0.028	0.0016	(0.0013,0.0016)	YES
KE-NC2	−0.052*	0.026	0.024	−0.052*	0.025	0.103	−0.0516	(-0.0516,-0.0516)	YES
GH-C	−0.017	0.017	0.003	−0.017	0.017	0.050	−0.0171	(-0.0171,-0.0171)	YES
KE-C	0.017	0.028	0.003	0.01	0.029	0.107	0.0066	(0.0066,0.0097)	YES
Employment Relations									
GH-NC	0.018	0.011	0.010	0.017	0.011	0.123	0.0171	(0.0171,0.0172)	YES
KE-NC1	−0.002	0.009	0.000	−0.001	0.009	0.064	−0.0012	(-0.0013,-0.0012)	YES
KE-NC2	−0.029*	0.012	0.040	−0.029**	0.011	0.156	−0.0294	(-0.0294,-0.0294)	YES
GH-C	0.006	0.005	0.004	0.006	0.005	0.101	0.0060	(0.006,0.006)	YES
KE-C	−0.021*	0.010	0.028	−0.019	0.011	0.094	−0.0174	(-0.0186,-0.0174)	YES
Workplace Safety									
GH-NC	0.039**	0.015	0.038	0.048***	0.013	0.297	0.0508	(0.0478,0.0508)	YES
KE-NC1	0.035*	0.014	0.032	0.035*	0.015	0.108	0.0350	(0.0349,0.035)	YES
KE-NC2	−0.016	0.040	0.006	−0.017	0.027	0.316	−0.0177	(-0.0177,-0.0174)	YES
GH-C	0.027**	0.009	0.032	0.026**	0.009	0.101	0.0261	(0.0261,0.0263)	YES
KE-C	0.014	0.014	0.007	0.045**	0.017	0.124	0.0553	(0.045,0.0553)	YES

(*continued on next page*)

(*continued*)

	Uncontrolled			Controlled					
	Effect	SE	R2	Effect	SE	R2	Beta	Bounds on the treatment effect (δ = 1, Rmax = 1.3*R)	Treatment effect excludes 0
Sustainability Management									
GH-NC	−0.024	0.021	0.004	−0.024	0.020	0.091	−0.0243	(-0.0243,-0.0242)	YES
KE-NC1	−0.013	0.021	0.002	−0.015	0.020	0.125	−0.0148	(-0.0148,-0.0145)	YES
KE-NC2	−0.017	0.026	0.003	−0.017	0.024	0.157	−0.0169	(-0.0169,-0.0169)	YES
GH-C	0.031	0.016	0.013	0.031*	0.015	0.132	0.0307	(0.0307,0.0307)	YES
KE-C	0.096***	0.022	0.104	0.101***	0.023	0.171	0.1072	(0.1012,0.1072)	YES
Conflict Resolution									
GH-NC	−0.045**	0.015	0.031	−0.448	0.015	0.098	−0.0445	(-0.0448,-0.0445)	YES
KE-NC1	0.032**	0.011	0.034	0.031**	0.011	0.057	0.0311	(0.0311,0.0313)	YES
KE-NC2	−0.018	0.022	0.005	−0.018	0.021	0.136	−0.0184	(-0.0184,-0.0184)	YES
GH-C	−0.007	0.008	0.003	−0.007	0.007	0.082	−0.0072	(-0.0072,-0.0071)	YES
KE-C	0.05**	0.015	0.070	0.07***	0.017	0.155	0.0820	(0.082,0.0701)	YES

[1] Significance of the effects based on p-values: * at a 10% level; ** at a 5% level; *** at a 1% level.
[2] Cases where the results differ between the "Controlled" and "Uncontrolled" regression are underlined.
[3] Bounds on the treatment effect are calculated using Oster (2019)'s Stata code psacalc. β is the effect with δ = 1 and Rmax = 1.3*R. Delta, δ, is a coefficient of proportionality that describes how large the effect of unobservables needs to be in proportion to the effect of observables for the treatment effect to be equal to 0, given a maximum value of the R-squared.

H. Effects of the initiatives at the Key indicator level

Appendix H. Table 1
Effects of the interventions on the key indicators for the "Soil Quality" (SQ), "Water Quality" (WQ) and "Species Diversity" (SD) sub-themes.

Case Study	Simple Comparison between Groups		Entropy	
	Mean Control	Mean Intervention	ATT	Robust Standard Error
SQ_1 Share of agricultural area on which insecticides are applied				
GH-NC	2%	2%	0%	0.009
KE-NC1	6%	4%	−1%	0.014
KE-NC2	38%	29%	−8%*	0.037
GH-C	79%	71%	−7%*	0.029
KE-C	56%	25%	−32%***	0.055
SQ_2 Share of agricultural area on which fungicides are applied				
GH-NC	1%	0%	0%	0.003
KE-NC1	3%	1%	−1%*	0.007
KE-NC2	35%	30%	−4%	0.039
GH-C	54%	41%	−15%**	0.047
KE-C	40%	48%	5%	0.048
SQ_3 Share of agricultural area on which herbicides are applied				
GH-NC	60%	70%	8%	0.044
KE-NC1	1%	0%	0%	0.005
KE-NC2	1%	0%	0%	0.003
GH-C	20%	16%	1%	0.034
KE-C	18%	7%	−17%***	0.046
SQ_5 Amount of Mineral N fertiliser applied (kg/ha)				
GH-NC	31.46	33.17	−2.08	1.368s
KE-NC1	28.93	32.25	0.94	7.120
KE-NC2	30.07	29.86	1.11	2.770
GH-C	18.00	17.62	0.10	0.617
KE-C	18.32	11.24	−23.69***	4.000
SQ_6 Share of farmers preparing and applyng compost				
GH-NC	6%	31%	22%***	0.053
KE-NC1	18%	31%	14%	0.036
KE-NC2	3%	0%	−2%	0.013
GH-C	–	–	–	–
KE-C	15%	26%	18%**	0.059
SQ_7 Share of farmers applying mulch				
GH-NC	26%	7%	−12%**	0.041
KE-NC1	16%	15%	1%	0.046
KE-NC2	2%	0%	−2%	0.021
GH-C	0%	1%	1%	0.007
KE-C	2%	6%	2%	0.035
SQ_8 Share of farmers correctly disposing of waste				
GH-NC	0%	0%	0%	–
KE-NC1	35%	47%	10%	0.066
KE-NC2	0%	0%	0%	–
GH-C	34%	32%	−4%	0.055
KE-C	1%	0%	0%	0.001
SQ_4 Share of farmers using active ingredients highly persistent in the soil				
GH-NC	1%	0%	0%	0.006
KE-NC1	15%	6%	6%	0.035
KE-NC2	20%	15%	5%	0.060
GH-C	84%	73%	7%	0.051

(*continued on next page*)

Appendix H. Table 1 (*continued*)

Case Study	Simple Comparison between Groups		Entropy	
	Mean Control	Mean Intervention	ATT	Robust Standard Error
KE-C	61%	79%	−15%*	0.069
WQ_1 Share of farmers using active ingredients highly toxic to aquatic organisms				
GH-NC	99%	97%	−3%	0.042
KE-NC1	90%	84%	−22%***	0.056
KE-NC2	47%	17%	−13%*	0.060
GH-C	95%	85%	−2%	0.016
KE-C	94%	81%	−11%**	0.043
WQ_2 Share of farmers using active ingredients highly persistent in the water				
GH-NC	1%	1%	0%	0.009
KE-NC1	43%	12%	−23%***	0.053
KE-NC2	87%	75%	−14%*	0.070
GH-C	10%	10%	−1%	0.041
KE-C	67%	27%	−51%***	0.082
WQ_3 Share of farms managing the risk of nutrients pollution				
GH-NC	84%	88%	2%	0.038
KE-NC1	46%	45%	−1%	0.066
KE-NC2	55%	43%	−8%	0.081
GH-C	94%	97%	6%	0.036
KE-C	12%	32%	17%**	0.062
WQ_4 Share of farmers managing riparian strips extensively				
GH-NC	31%	20%	−12%	0.111
KE-NC1	46%	46%	−1%	0.073
KE-NC2	71%	78%	9%	0.130
GH-C	31%	45%	13%	0.146
KE-C	54%	71%	17%*	0.084
SD_1 Share of farmers using active ingredients toxic to bees				
GH-NC	87%	81%	3%	0.049
KE-NC1	26%	12%	10%	0.052
KE-NC2	92%	71%	21%*	0.086
GH-C	98%	96%	2%	0.017
KE-C	88%	75%	10%	0.090

Significance of the effects based on p-values: * at a 10% level; ** at a 5% level; *** at a 1% level.

Appendix H. Table 2

Effects of the interventions on the key indicators for the "Stability of Production" (SP), "Stability of Market" (SM), and "Value Creation" (VC) sub-themes.

Case Study	Simple Comparison between Groups		Entropy	
	Mean Control	Mean Intervention	ATT	Robust Standard Error
SP_1 Share of farmers affected by yield loss due to pests or diseases				
GH-NC	58%	77%	21%***	0.057
KE-NC1	82%	80%	−6%	0.062
KE-NC2	61%	52%	−8%	0.081
GH-C	35%	37%	4%	0.055
KE-C	45%	82%	47%***	0.068
SP_2 Share of farmers affected by yield loss due to lack of access to water				
GH-NC	83%	62%	−15%**	0.055
KE-NC1	69%	75%	5%	0.059
KE-NC2	4%	6%	3%	0.033
GH-C	34%	42%	8%	0.057
KE-C	13%	22%	12%*	0.058
SP_3 Share of farmers growing crop resistant varieties				
GH-NC	12%	39%	22%***	0.052
KE-NC1	79%	64%	−14%*	0.061
KE-NC2	26%	17%	−5%	0.061
GH-C	38%	37%	−12%*	0.059
KE-C	36%	18%	−24%**	0.073
SP_4 Share of farmers that experienced staff shortages				
GH-NC	0%	0%	0%	–
KE-NC1	9%	4%	−6%	0.043
KE-NC2	14%	8%	−8%	0.059
GH-C	17%	25%	8%	0.048
KE-C	25%	44%	24%***	0.070
SM_1 Average number of buyers per farm				
GH-NC	1.45	1.51	0.1	0.057
KE-NC1	2.16	2.24	0.11	0.134
KE-NC2	2.00	2.59	0.49	0.546
GH-C	1.52	1.57	−0.17	0.065
KE-C	3.76	4.70	1.14***	0.001
SM_2 Share of income generate by the main buyer				
GH-NC	35%	34%	−4%	0.032

(*continued on next page*)

Appendix H. Table 2 (*continued*)

Case Study	Simple Comparison between Groups		Entropy	
	Mean Control	Mean Intervention	ATT	Robust Standard Error
KE-NC1	56%	59%	−4%	0.033
KE-NC2	60%	68%	−3%	0.058
GH-C	19%	28%	0%	0.046
KE-C	53%	48%	4%	0.034
SM_3 Share of farmer having access to alternative market s				
GH-NC	99%	99%	1%	0.014
KE-NC1	34%	25%	−7%	0.072
KE-NC2	75%	76%	4%	0.072
GH-C	100%	99%	−1%	0.005
KE-C	45%	74%	33%***	0.074
SM_4 Share of farmers having a form of cooperation with buyers				
GH-NC	73%	60%	−13%*	0.053
KE-NC1	42%	65%	23%***	0.064
KE-NC2	26%	27%	2%	0.074
GH-C	73%	61%	−12%*	0.054
KE-C	60%	62%	2%	0.078
VC_1 Average weekly working hours/ha				
GH-NC	31.46	33.17	0.74	0.73
KE-NC1	28.93	32.25	2.17	1.72
KE-NC2	30.07	29.86	1.42	1.29
GH-C	18.00	17.62	−1.08	1.03
KE-C	18.32	11.24	−7.80***	1.43
VC_2 Share of farms on which new jobs were created				
GH-NC	0%	0%		
KE-NC1	9%	7%	−2%	0.032
KE-NC2	7%	4%	1%	0.028
GH-C	9%	10%	0%	0.037
KE-C	10%	26%	20%***	0.053
VC_3 Proportion of the externally sourced inputs that is locally produced				
GH-NC	13%	11%	−2%	0.025
KE-NC1	56%	56%	1%	0.028
KE-NC2	27%	25%	−3%	0.041
GH-C	25%	30%	5%	0.031
KE-C	10%	11%	1%	0.024

Significance of the effects based on p-values: * at a 10% level; ** at a 5% level; *** at a 1% level.

Appendix H. Table 3

Effects of the interventions on the key indicators for the “Capacity Development” (CD), “Employment Relations” (ER), and “Workplace Safety” (WS) sub-themes.

Case Study	Simple Comparison between Groups		Entropy	
	Mean Control	Mean Intervention	ATT	Robust Standard Error
CD_1 Share of farmers that received at least 1 day of training/year				
GH-NC	42%	65%	17%**	0.058
KE-NC1	67%	77%	7%	0.057
KE-NC2	53%	33%	−22%**	0.078
GH-C	26%	31%	−9%	0.058
KE-C	51%	91%	39%***	0.067
CD_2 Share of farmers having adequate access to extension services				
GH-NC	23%	15%	−3%	0.043
KE-NC1	15%	16%	−2%	0.053
KE-NC2	20%	9%	−9%	0.051
GH-C	18%	18%	2%	0.044
KE-C	27%	14%	−6%	0.056
ER_1 Share of farmers providing written contracts to workers				
GH-NC	0%	0%	0%	–
KE-NC1	0%	0%	0%	–
KE-NC2	4%	3%	0%	0.030
GH-C	0%	0%	0%	–
KE-C	1%	0%	0%	–
ER_2 Share of farmers having or providing social protection to workers				
GH-NC	0%	6%	6%	0.063
KE-NC1	29%	25%	−6%	0.078
KE-NC2	7%	6%	1%	0.048
GH-C	20%	26%	4%	0.070
KE-C	30%	19%	−3%	0.076
ER_3 Share of farmers paying more than the regional minimum wages				
GH-NC	6%	22%	16%***	0.039
KE-NC1	9%	6%	−6%	0.049
KE-NC2	14%	5%	−8%	0.049
GH-C	4%	7%	4%	0.023

(*continued on next page*)

Appendix H. Table 3 (*continued*)

Case Study	Simple Comparison between Groups		Entropy	
	Mean Control	Mean Intervention	ATT	Robust Standard Error
KE-C	16%	26%	5%	0.073
ER_3 Share of farmers allowing regular breaks				
GH-NC	98%	100%	0%	0.007
KE-NC1	76%	87%	12%*	0.063
KE-NC2	93%	92%	3%	0.056
GH-C	100%	99%	−1%	0.005
KE-C	86%	68%	−14%*	0.070
ER_5 Share of farms on which the risk of child performing hazarduous work was considered high				
GH-NCS	44%	33%	−7%	0.057
KE-NC1	0%	0%	0%	–
KE-NC2	1%	2%	1%	0.021
GH-C	0%	1%	0%	0.009
KE-C	0%	2%	2%	0.018
ER_6 Share of farms on which the risk of child labor affecting school performance was considered high				
GH-NC	47%	33%	−10%	0.057
KE-NC1	1%	0%	−2%	0.016
KE-NC2	1%	6%	4%	0.033
GH-C	1%	1%	0%	0.011
KE-C	1%	1%	1%	0.013
WS_1 Share of farmers using pesticides classified as acute toxic to the health				
GH-NC	44%	34%	−15%*	0.063
KE-NC1	28%	10%	−22%***	0.055
KE-NC2	74%	61%	−13%*	0.060
GH-C	74%	72%	−3%	0.018
KE-C	80%	65%	−11%**	0.043
WS_2 Share of farmers using pesticides classified as acutely toxic when inhaled				
GH-NC	98%	96%	−3%	0.042
KE-NC1	85%	66%	−20%***	0.056
KE-NC2	43%	13%	−13%*	0.059
GH-C	95%	85%	−2%	0.016
KE-C	94%	81%	−11%**	0.043
WS_3 Share of farmers using pesticides having adverse long term effects				
GH-NC	34%	15%	−21%***	0.053
KE-NC1	34%	10%	−16%***	0.046
KE-NC2	80%	71%	−6%	0.074
GH-C	70%	63%	−8%	0.058
KE-C	60%	13%	−51%***	0.064
WS_4 Share of farmers using or providing adequate protective gear				
GH-NC	33%	47%	9%	0.052
KE-NC1	82%	86%	4%	0.050
KE-NC2	55%	43%	−4%	0.073
GH-C	46%	51%	3%	0.045
KE-C	58%	59%	5%	0.076

Significance of the effects based on p-values: * at a 10% level; ** at a 5% level; *** at a 1% level.

Appendix H. Table 4
Effects of the interventions on the key indicators for the "Sustainability Management Plan" (MP) and "Conflict Resolution" (CR) sub-themes.

Case Study	Simple Comparison between Groups		Entropy	
	Mean Control	Mean Intervention	ATT	Robust Standard Error
MP_1 Share of farmers that recognize sustainability principles				
GH-NC	45%	51%	3%	0.059
KE-NC1	57%	51%	−7%	0.065
KE-NC2	36%	26%	−2%	0.070
GH-C	78%	83%	−1%	0.040
KE-C	45%	67%	23%**	0.077
MP_2 Share of farmers having sustainability measures planned				
GH-NC	48%	51%	0%	0.059
KE-NC1	50%	65%	13%*	0.064
KE-NC2	47%	26%	−19%*	0.074
GH-C	68%	71%	−1%	0.052
KE-C	75%	96%	24%***	0.059
MP_4 Share of farmers aware and informed about future policy changes				
GH-NC	56%	24%	−25%***	0.056
KE-NC1	69%	75%	3%	0.058
KE-NC2	5%	7%	4%	0.038
GH-C	28%	29%	−4%	0.056
KE-C	29%	56%	20%*	0.081
MP_3 Share of farmers aware and informed about future market challenges				
GH-NC	47%	49%	−3%	0.059
KE-NC1	79%	82%	1%	0.052

(*continued on next page*)

Appendix H. Table 4 (*continued*)

Case Study	Simple Comparison between Groups		Entropy	
	Mean Control	Mean Intervention	ATT	Robust Standard Error
KE-NC2	46%	41%	−2%	0.080
GH-C	41%	51%	9%	0.059
KE-C	28%	82%	53%***	0.066
MP_5 Share of farmers aware and informed about climate change challenges				
GH-NC	56%	54%	−3%	0.058
KE-NC1	50%	36%	−14%*	0.066
KE-NC2	62%	54%	−4%	0.082
GH-C	16%	26%	12%*	0.044
KE-C	39%	54%	12%	0.078
CR_1 Share of farmers that cooperated with other farmers				
GH-NC	84%	81%	−7%*	0.037
KE-NC1	27%	42%	16%**	0.053
KE-NC2	44%	40%	−3%	0.061
GH-C	79%	82%	2%	0.044
KE-C	20%	24%	5%	0.055
CR_3 Share of farmers that had conflicts over water quality				
GH-NC	1%	0%	−1%	0.012
KE-NC1	0%	1%	1%	0.009
KE-NC2	5%	2%	−2%	0.024
GH-C	1%	0%	−2%	0.013
KE-C	0%	2%	2%	0.017
CR_2 Share of farmers that had conflicts over access to water				
GH-NC	11%	3%	−8%*	0.034
KE-NC1	0%	1%	1%	0.009
KE-NC2	16%	22%	9%	0.063
GH-C	0%	1%	0%	0.009
KE-C	4%	7%	0%	0.045

Significance of the effects based on p-values: * at a 10% level; ** at a 5% level; *** at a 1% level.

References

Adamtey, N., Musyoka, M.W., Zundel, C., Cobo, J.G., Karanja, E., Fiaboe, K.K.M., Muriuki, A., Mucheru-Muna, M., Vanlauwe, B., Berset, E., Messmer, M.M., Gattinger, A., Bhullar, G.S., Cadisch, G., Fliessbach, A., Mäder, P., Niggli, U., Foster, D., 2016. Productivity, profitability and partial nutrient balance in maize-based conventional and organic farming systems in Kenya. Agric. Ecosyst. Environ. 235, 61–79. https://doi.org/10.1016/j.agee.2016.10.001.

Akoyi, K.T., Maertens, M., 2018. Walk the talk: private sustainability standards in the Ugandan coffee sector. J. Dev. Stud. 54, 1792–1818. https://doi.org/10.1080/00220388.2017.1327663.

Anderberg, S., 2020. The contribution of organic agriculture to poverty reduction. In: Breiling, M., Anbumozhi, V. (Eds.), Vulnerability of Agricultural Production Networks and Global Food Value Chains Due to Natural Disasters. Economic Research Institute for ASEAN and East Asia, Jakarta, Indonesia, pp. 42–72.

Ansah, E.O., Kaplowitz, M.D., Lupi, F., Kerr, J., 2020. Smallholder participation and procedural compliance with sustainable cocoa certification programs. Agroecol. Sustain. Food Syst. 44, 54–87. https://doi.org/10.1080/21683565.2019.1579776.

Arulnathan, V., Heidari, M.D., Doyon, M., Li, E., Pelletier, N., 2020. Farm-level decision support tools: a review of methodological choices and their consistency with principles of sustainability assessment. J. Clean. Prod. 256, 120410 https://doi.org/10.1016/j.jclepro.2020.120410.

Bachmann, F., 2012. Potential and limitations of organic and fair trade cotton for improving livelihoods of smallholders: evidence from Central Asia. Renew. Agric. Food Syst. 27, 138–147. https://doi.org/10.1017/S1742170511000202.

Bennett, M., Franzel, S., 2013. Can organic and resource-conserving agriculture improve livelihoods? A synthesis. Int. J. Agric. Sustain. 11, 193–215. https://doi.org/10.1080/14735903.2012.724925.

Bolwig, S., Gibbon, P., Jones, S., 2009. The economics of smallholder organic contract farming in tropical Africa. World Dev. 37, 1094–1104. https://doi.org/10.1016/j.worlddev.2008.09.012.

Bonisoli, L., Galdeano-Gómez, E., Piedra-Muñoz, L., Pérez-Mesa, J.C., 2019. Benchmarking agri-food sustainability certifications: evidences from applying SAFA in the Ecuadorian banana agri-system. J. Clean. Prod. 236, 117579 https://doi.org/10.1016/j.jclepro.2019.07.054.

Breusch, T.S., Pagan, A.R., 1979. A simple test for heteroscedasticity and random coefficient variation. Econometrica 47, 1287. https://doi.org/10.2307/1911963.

Bryan, M.L., Rice, N., Roberts, J., Sechel, C., 2020. Mental Health and Employment : a Bounding Approach Using Panel Data 54.

Chiputwa, B., Spielman, D.J., Qaim, M., 2015. Food standards, certification, and poverty among coffee farmers in Uganda. World Dev. 66, 400–412. https://doi.org/10.1016/j.worlddev.2014.09.006.

Curran, M., Lazzarini, G., Baumgart, L., Gabel, V., Blockeel, J., Epple, R., Stolze, M., Schader, C., 2020. Representative farm-based sustainability assessment of the organic sector in Switzerland using the SMART-farm tool. Front. Sustain. Food Syst. 4, 554362 https://doi.org/10.3389/fsufs.2020.554362.

De Fries, R.S., Fanzo, J., Mondal, P., Remans, R., Wood, S.A., 2017. Is voluntary certification of tropical agricultural commodities achieving sustainability goals for small-scale producers? A review of the evidence. Environ. Res. Lett. 12, 033001 https://doi.org/10.1088/1748-9326/aa625e.

Dietz, T., Grabs, J., Chong, A.E., 2019. Mainstreamed Voluntary Sustainability Standards and Their Effectiveness: Evidence from the Honduran Coffee Sector: Effectiveness of Mainstreamed VSS. Regul. Gov. https://doi.org/10.1111/rego.12239.

FAO, WHO, 2007. In: Organically Produced Foods, third ed. Codex alimentarius. World Health Organization : Food and Agriculture Organization of the United Nations, Rome.

Gibbon, P., Lin, Y., Jones, S., 2009. Revenue Effects of Participation in Smallholder Organic Cocoa Production in Tropical Africa a Case Study. Danish Institute for International Studies, Copenhagen.

Gibbon, P., Ponte, S., Lazaro, E. (Eds.), 2010. Global Agro-Food Trade and Standards. Palgrave Macmillan UK, London. https://doi.org/10.1057/9780230281356.

Haggblade, S., Minten, B., Pray, C., Reardon, T., Zilberman, D., 2017. The herbicide revolution in developing countries: patterns, causes, and implications. Eur. J. Dev. Res. 29, 533–559. https://doi.org/10.1057/s41287-017-0090-7.

Hainmueller, J., 2012. Entropy balancing for causal effects: a multivariate reweighting method to produce balanced samples in observational studies. Polit. Anal. 20, 25–46. https://doi.org/10.1093/pan/mpr025.

Hainmueller, J., Xu, Y., 2013. Ebalance: a Stata package for entropy balancing. J. Stat. Software 18.

IFAD, 2013. Smallholders, Food Security, and the Environment.

IPES-Food Report, 2020. Biovision Foundation for Ecological Development & IPES-Food. 2020. Money Flows: what Is Holding Back Investment in Agroecological Research for Africa? Biovision Foundation for Ecological Development & International Panel of Experts on Sustainable Food Systems.

ITC, 2016. Social and Environmental Standards: Contributing to More Sustainable Value Chains. ITC/EUI, Geneva.

Jena, P.R., Chichaibelu, B.B., Stellmacher, T., Grote, U., 2012, 2012 The impact of coffee certification on small-scale producers' livelihoods: a case study from the Jimma Zone, Ethiopia: P. R. Jena et al./Agricultural Economics 43, 427–438. https://doi.org/10.1111/j.1574-0862.2012.00594.x. Agric. Econ. 43, 429–440.

Jouzi, Z., Azadi, H., Taheri, F., Zarafshani, K., Gebrehiwot, K., Van Passel, S., Lebailly, P., 2017. Organic farming and small-scale farmers: main opportunities and challenges. Ecol. Econ. 132, 144–154. https://doi.org/10.1016/j.ecolecon.2016.10.016.

Kamau, J.W., Schader, C., Biber-Freudenberger, L., Stellmacher, T., Amudavi, D.M., Landert, J., Blockeel, J., Whitney, C., Borgemeister, C., 2021. A holistic sustainability assessment of organic (certified and non-certified) and non-organic smallholder farms in Kenya. Environ. Dev. Sustain. https://doi.org/10.1007/s10668-021-01736-y.

Kamau, J.W., Stellmacher, T., Biber-Freudenberger, L., Borgemeister, C., 2018. Organic and conventional agriculture in Kenya: a typology of smallholder farms in Kajiado and Murang'a counties. J. Rural Stud. 57, 171–185. https://doi.org/10.1016/j.jrurstud.2017.12.014.

Kilcher, L., 2007. How Organic Agriculture Contributes to Sustainable Development.

Landert, J., Pfeifer, C., Carolus, J.F., 2020. Assessing agro-ecological practices using a combination of three sustainability assessment tools. Landbauforsch. J. Sustain. Org. Agric. Syst. 129–144. https://doi.org/10.3220/LBF1612794225000.
Lotter, D., 2015. Facing food insecurity in Africa: why, after 30 years of work in organic agriculture, I am promoting the use of synthetic fertilizers and herbicides in small-scale staple crop production. Agric. Hum. Val. 32, 111–118. https://doi.org/10.1007/s10460-014-9547-x.
Meemken, E.-M., 2020. Do smallholder farmers benefit from sustainability standards? A systematic review and meta-analysis. Global Food Secur. 26, 100373 https://doi.org/10.1016/j.gfs.2020.100373.
Meemken, E.-M., Qaim, M., 2018a. Can private food standards promote gender equality in the small farm sector? J. Rural Stud. 58, 39–51. https://doi.org/10.1016/j.jrurstud.2017.12.030.
Meemken, E.-M., Qaim, M., 2018b. Organic agriculture, food security, and the environment. Annu. Rev. Resour. Econ. 10, 39–63. https://doi.org/10.1146/annurev-resource-100517-023252.
Meinshausen, F., Richter, T., Blockeel, J., Huber, B., 2019. Internal Control Systems in Organic Agriculture: Significance, Opportunities and Challenges.
NEPAD, 2013. Agriculture in Africa: transformation and outlook, p. 25.
Oberč, B.P., 2020. Approaches to Sustainable Agriculture: Exploring the Pathways towards the Future of Farming. IUCN, International Union for Conservation of Nature. https://doi.org/10.2305/IUCN.CH.2020.07.en.
Oster, E., 2019. Unobservable selection and coefficient stability: theory and evidence. J. Bus. Econ. Stat. 37, 187–204. https://doi.org/10.1080/07350015.2016.1227711.
Oya, C., Schaefer, F., Skalidou, D., McCosker, C., Langer, L., 2017. Effects of certification schemes for agricultural production on socio-economic outcomes in low- and middle-income countries: a systematic review. Campbell Syst. Rev. 13, 1–346. https://doi.org/10.4073/csr.2017.3.
Parrot et al., 2003. Organic and Like-Minded Movements in Africa.
Parrott, N., Ssekyewa, C., Makunike, C., Ntambi, S.M., 2006. 12 Organic Farming in Africa.
Pintér, L., Hardi, P., Martinuzzi, A., Hall, J., 2012. Bellagio STAMP: principles for sustainability assessment and measurement. Ecol. Indicat. 17, 20–28. https://doi.org/10.1016/j.ecolind.2011.07.001.
Rosina Bara, C., Jarquin Gálvez, R., Reyes Hernández, H., Fortanelli Martínez, J., 2018. Adaptation of a participatory organic certification system to the organic products law in six local markets in Mexico. Agroecol. Sustain. Food Syst. 42, 48–76. https://doi.org/10.1080/21683565.2017.1359736.
Rosinger, C., 2013. Organic Agriculture and Certification in Eastern Africa 8.
Ruben, R., 2015. Coffee certification in east Africa: impact on farmers, families and cooperatives. Mt. Res. Dev. 35, 306. https://doi.org/10.1659/mrd.mm164.
Schader, C., Baumgart, L., Landert, J., Muller, A., Ssebunya, B., Blockeel, J., Weisshaidinger, R., Petrasek, R., Mészáros, D., Padel, S., Gerrard, C., Smith, L., Lindenthal, T., Niggli, U., Stolze, M., 2016a. Using the sustainability monitoring and assessment routine (SMART) for the systematic analysis of trade-offs and synergies between sustainability dimensions and themes at farm level. Sustainability 8, 274. https://doi.org/10.3390/su8030274.
Schader, C., Baumgart, L., Landert, J., Muller, A., Ssebunya, B., Blockeel, J., Weisshaidinger, R., Petrasek, R., Mészáros, D., Padel, S., Gerrard, C., Smith, L., Lindenthal, T., Niggli, U., Stolze, M., 2016b. Using the sustainability monitoring and assessment routine (SMART) for the systematic analysis of trade-offs and synergies between sustainability dimensions and themes at farm level. Sustainability 8, 274. https://doi.org/10.3390/su8030274.
Schader, C., Curran, M., Heidenreich, A., Landert, J., Blockeel, J., Baumgart, L., Ssebunya, B., Moakes, S., Marton, S., Lazzarini, G., Niggli, U., Stolze, M., 2019. Accounting for uncertainty in multi-criteria sustainability assessments at the farm level: improving the robustness of the SMART-Farm Tool. Ecol. Indicat. 106, 105503 https://doi.org/10.1016/j.ecolind.2019.105503.
Schader, C., Heidenreich, A., Kadzere, I., Egyir, I., Muriuki, A., Bandanaa, J., Clottey, J., Ndungu, J., Grovermann, C., Lazzarini, G., Blockeel, J., Borgemeister, C., Muller, A., Kabi, F., Fiaboe, K., Adamtey, N., Huber, B., Niggli, U., Stolze, M., 2021. How is organic farming performing agronomically and economically in sub-Saharan Africa? Global Environ. Change 102325. https://doi.org/10.1016/j.gloenvcha.2021.102325.
Scialabba, N., 2002. Organic agriculture, environment and food security. Environment and Natural Resources Series. Food and Agriculture Organization of the United Nations (FAO, Rome.
Sheahan, M., Barrett, C.B., 2017. Ten striking facts about agricultural input use in Sub-Saharan Africa. Food Pol. 67, 12–25. https://doi.org/10.1016/j.foodpol.2016.09.010.
Siebrecht, N., 2020. Sustainable agriculture and its implementation gap—overcoming obstacles to implementation. Sustainability 12, 3853. https://doi.org/10.3390/su12093853.
Ssebunya, B.R., Schader, C., Baumgart, L., Landert, J., Altenbuchner, C., Schmid, E., Stolze, M., 2019. Sustainability performance of certified and non-certified smallholder coffee farms in Uganda. Ecol. Econ. 156, 35–47. https://doi.org/10.1016/j.ecolecon.2018.09.004.
Stewart, R., Langer, L., Da Silva, N.R., Muchiri, E., Zaranyika, H., Erasmus, Y., Randall, N., Rafferty, S., Korth, M., Madinga, N., Wet, T., 2015. The effects of training, innovation and new technology on african smallholder farmers' economic outcomes and food security: a systematic review. Campbell Syst. Rev. 11, 1–224. https://doi.org/10.4073/csr.2015.16.
Tey, Y.S., Li, E., Bruwer, J., Abdullah, A.M., Brindal, M., Radam, A., Ismail, M.M., Darham, S., 2017. Factors influencing the adoption of sustainable agricultural practices in developing countries: a review. Environ. Eng. Manag. J. 16, 337–349. https://doi.org/10.30638/eemj.2017.034.
Tripp, R., 2001. Agricultural technology policies for rural development. Dev. Pol. Rev. 19, 479–489. https://doi.org/10.1111/1467-7679.00146.
Twarog, S., 2006. Organic agriculture: a trade and sustainable development opportunity for developing countries. In: Trade and Environment Review 2006. United Nations Conference on Trade and Development (UNCTAD) Trade and Environment Review, UN, pp. 141–223. https://doi.org/10.18356/6ad7c431-en.
UNEP-UNCTAD, 2008. Organic Agriculture and Food Security in Africa. UNEP-UNCTAD. 2008. UNEP-UNCTAD, Geneva, Switzerland, p. 47. Available at: Web site. http://www.unep-unctad.org/cbtf.
van der Mheen-Sluijer, J., Innocenti, E.D., 2011. Socio-economic Impact of Organic Farming in East Africa. Van der Mheen-Sluijer, 2011.
White, H., 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica 48, 817. https://doi.org/10.2307/1912934.
Willer, E.H., Lernoud, J., 2019. The World of Organic Agriculture - Statistics and Emerging Trends 2019 356.
Winter, E., Marton, S.M.R.R., Baumgart, L., Curran, M., Stolze, M., Schader, C., 2020. Evaluating the sustainability performance of typical conventional and certified coffee production systems in Brazil and Ethiopia based on expert judgements. Front. Sustain. Food Syst. 4, 49. https://doi.org/10.3389/fsufs.2020.00049.