Impact of participatory guarantee systems on sustainability outcomes: the case of vegetable farming in Vietnam

Christian Grovermann, Pham Van Hoi, Nguyen Thi Bich Yen, Pepijn Schreinemachers, Minh Ngo Hai & Pierre Ferrand

Published online: 09 Apr 2024.

Submit your article to this journal ↗

Article views: 185

View related articles ↗

View Crossmark data ↗

Routledge
Taylor & Francis Group

Impact of participatory guarantee systems on sustainability outcomes: the case of vegetable farming in Vietnam

Christian Grovermann[a], Pham Van Hoi ⬤[b], Nguyen Thi Bich Yen[c], Pepijn Schreinemachers[d], Minh Ngo Hai[e] and Pierre Ferrand[f]

[a]Research Institute of Organic Agriculture, Frick, Switzerland; [b]Center for Agricultural Research and Ecological Studies, Vietnam National University of Agriculture, Hanoi, Vietnam; [c]Faculty of Natural Resources and Environment, Vietnam National University of Agriculture, Hanoi, Vietnam; [d]World Vegetable Center, Bangkok, Thailand; [e]Faculty of Economics and Rural Development, Vietnam National University of Agriculture, Hanoi, Vietnam; [f]Food and Agriculture Organization of the United Nations, Rome, Italy

ABSTRACT

Sustainable food systems require mechanisms that assure consumers about the sustainability of agricultural production. Building on the existing literature on the impact of sustainability standards, this study is the first to assess the effectiveness of participatory guarantee systems (PGS) for the certification of organic produce. The study uses representative farm-level data on local vegetable value chains in northern Vietnam and uses a broad set of sustainability outcomes as well as counterfactual analysis, including systematic robustness checks. The results show that PGS significantly improves farm profitability (+117%), agroecology performance (+40%), and gives farmers more choice of sales channels (+23%). However, PGS had no significant effect on returns to labour and reduced the average crop yield. Capacity development on nursery practices, transplanting of healthy seedlings rather than direct seeding, reduced tillage, and collective crop planning and management are some of the innovations that can counter adverse effects on crop yields, increase soil health as well as improve returns to labour, and thus attract more youth to farming. Overall, the study shows that organic PGS can make vegetable production more economically viable and more agroecologically sustainable.

1. Introduction

It has been estimated that the external health and environmental costs of global food systems are double the value of all food sold (von Braun & Hendriks, 2023). At the same time, environmental degradation increases the risk of food shortages, especially in low- and middle-income countries. There is a widespread consensus among researchers and policymakers that healthier, more sustainable and more inclusive food systems are needed to tackle critical problems, such as soil degradation, water scarcity, reduced ecosystem services, climate change, economic inequality, food insecurity and unhealthy diets UNSG, 2020). Agroecology has evolved from a science of sustainable agriculture to a holistic approach to the redesign of agricultural and food systems based on community empowerment, diversity and systems thinking (Altieri, 1995; IPES-Food, 2016; Gliessman, 2018). It is now increasingly recognized by experts, policymakers and the public as a promising strategy to profoundly change the way food is produced (CFS, 2021; FAO, 2018a; HLPE, 2019). The concept of agroecology goes beyond a set of sustainable production practices by

including aspects of knowledge sharing, social equity, and circularity, while addressing issues of environmental integrity, food and nutrition security, and livelihoods (FAO, 2019; Kerr et al., 2021; van der Ploeg et al., 2019; Wezel et al., 2020). Agroecology therefore requires embedded solutions that systematically combine technical, social and institutional innovations addressing key bottlenecks in the food system.

Participatory guarantee systems (PGS) are one such approach, which supports the diffusion of agroecology, especially in low- and middle-income countries where third-party certification systems are less developed and expensive. PGS promotes fair prices, reduces dependence on external inputs, and strengthens interactions between buyers, growers, and other stakeholders (Home et al., 2017; Loconto and Hakanata, 2018). Most PGS schemes are certified based on the principles of organic agriculture (IFOAM, 2019), but the approach can also be used for other types of farming. Based on trust and monitoring through social and business networks, PGS offers a locally focused lower-cost alternative to third-party organic certification, while providing quality assurance in local value chains (Home et al., 2017; Loconto and Hakanata, 2018). In 2022, 323 PGS initiatives were recorded in 76 countries globally, with around 1.3 million producers and 1 million hectares of land certified (Willer et al., 2023).

Southeast Asia, just like many other regions, has undergone important changes in its food systems in recent decades, mostly due to rapid economic growth, increasing population and demographic shifts, rapid urbanization, changing food consumption behaviour and climate change (FAO, 2018b). The region is not only facing biodiversity loss and climate change impacts, but also the burden of unsafe food and malnutrition. There is a growing distrust of consumers towards food quality and food safety and increasing threats to the health of farm families, both caused by high and rising levels of agricultural pesticide use (Ha et al., 2019; Schreinemachers et al., 2015). Problems are particularly acute in high-value and pest-vulnerable crops such as fruit and vegetables (Praneetvatakul et al., 2013; Wanwimolruk et al., 2016). It has been estimated that about 75% of pesticides used in vegetable production in Southeast Asia are in excess of the economic optimum (Grovermann et al., 2013; Schreinemachers et al., 2020). In response to the chronic overuse and misuse of pesticides in agriculture, governments in Southeast Asia have sought to improve food safety by promoting integrated pest management (IPM) alongside standards of good agricultural practices (GAP). These standards help to assure the quality of farm produce by defining how food should be produced and handled. However, GAP standards may not give farmers a strong enough incentive to adopt more sustainable crop management methods (Amekawa, 2013; Schreinemachers et al., 2012). Organic certification aims to eliminate the use of agrochemicals in agricultural production altogether and organic labels for domestic markets are now common across Southeast Asia.

The impacts of sustainability standards, including organic vegetable certification, have been widely analysed. Economic effects have generally received the most attention in the literature (Bolwig et al., 2009; Crowder & Reganold, 2015; Grovermann et al., 2021; Mendoza, 2004; Schleifer & Sun, 2020: Ssebunya et al., 2019; Tran & Goto, 2019). Others have studied the effects on the adoption of good agricultural practices, gender equality, or pesticide use (Ibanez & Blackman, 2016; Meemken & Qaim, 2018a; Schreinemachers et al., 2012). Very few studies have quantified the effect on agroecological outcomes, one notable exception being Blockeel et al. (2023). We address this gap by combining various economic and agroecology performance indicators in a holistic evaluation of the impact of organic PGS adoption among vegetable farmers in Vietnam.

In the context of low- and middle-income countries, most previous studies have looked into the impact of labels for export crops such as coffee and cocoa, while there is little evidence for the impact of labels for food sold in local markets. This study therefore focuses on organic PGS certification of local vegetable value chains. The Vietnamese public third-party organic standard (TCVN 11041-1 issued in 2017) exists alongside various international third-party standards (e.g. 'USDA Organic' and 'Organic JAS') as well as numerous private organic PGS standards. Established in 2008, PGS-Vietnam is the first and most notable PGS standard and documenting its impact has clear policy relevance. The national food system action plan of Vietnam specifies a minimum target of 2.5% of the agricultural area for organic production by 2030 as well as a doubling of organic fertilizer from 2020 levels (USDA, 2023).

Value-chain actors and policymakers need a better understanding of the socioeconomic and ecological

impact of adopting organic PGS by vegetable farmers. We aim to determine the causal effect of farmers' participation in PGS-Vietnam on a broad set of sustainability outcomes. We used a representative sample of vegetable farms and employed entropy balancing combined with regression adjustment for counterfactual analysis. We postulate that PGS certification delivers economic, environmental, and social benefits, inducing a shift towards more viable and agroecological vegetable production. To our knowledge, this is the first impact study on organic PGS in Vietnam and elsewhere. We investigate the sensitivity of our findings to model assumptions and use alternative model specifications to identify the underlying mechanisms and to improve the accuracy of our results. The analysis finds that PGS certification can be instrumental in achieving sustainability goals, but adoption is constrained by lower yields and high labour requirements. The next section explains the data sources and econometric methods used to assess impact. Impact estimates are then presented in Section 3 and discussed in Section 4. The final section derives policy implications from the results.

2. Material and methods

2.1. Analytical framework

Causal attribution in impact evaluation studies involves determining the difference between two potential outcomes for the same unit, one is observed, and the other is unobserved (Imbens & Rubin, 2015; Imbens & Wooldridge, 2009). The observed outcome is what actually occurred, while the unobserved outcome is what would have occurred had no or a different intervention been assigned, i.e. the counterfactual. In the case of a binary treatment, the two potential outcomes for each individual i are denoted as y_i^1 for state 1 (intervention) and y_i^0 for state 0 (counterfactual). In reality, data for each individual is only available for one state, while it is missing for the other state. This requires the careful construction of the counterfactual using a control group. Consequently, mean outcome differences between the intervention group and the control group are used to estimate average treatment effects rather than calculating individual treatment effects (Schafer & Kang, 2008). Other than in randomized experiments self-selection bias is a key challenge for causal inference in many ex-post impact evaluations, due to non-random assignment of the intervention (Imbens &

Wooldridge, 2009). In our case, self-selection bias stems from the fact that producers who choose to participate in the standard may be different from producers who do not participate even before they adopted it (Meemken & Qaim, 2018a; Ssebunya et al., 2019). It was therefore necessary to process the data so that producers in the intervention and comparison groups no longer significantly differed in observable characteristics that are unrelated to certification, but that may affect the outcomes. Propensity score matching and reweighting are widely used techniques in ex-post impact evaluation to adjust for differences in known confounders and mitigate self-selection bias (Austin, 2011; Hirano et al., 2003; Knook et al., 2018; Lopez-Avila et al., 2017). Unlike instrumental variable (IV) regression and other regression-based approaches, matching and reweighting do not assume a normal distribution of the residuals. However, these methods only correct selection bias for the observed confounding variables (Hirano et al., 2003; White & Raitzer, 2017), not the unobserved covariates, such as farmer self-motivation for instance. It is however difficult to find one or more appropriate instruments[1] that work for multiple separate IV estimations as there are nine outcome variables in this study. Therefore, this study relies on inverse probability and entropy weights (Austin, 2011; Hainmueller, 2012). We used these two alternative model specifications to verify the robustness of results. We also tested the sensitivity of our estimates to omitted variable bias.

2.2. Identification of counterfactual and data collection

Data for this study come from a survey of peri-urban vegetable farmers who produced white cabbage from August 2022 to March 2023 with data collected from March to April 2023. White cabbage was chosen as a focus crop as it is widely grown, commercially important and representative of the local vegetable cropping system. The survey covered 9 districts of three Vietnamese provinces (Hoabinh, Hanam, Hanoi). The organic PGS-Vietnam scheme is active in this peri-urban area near the capital Hanoi and farmers supply vegetables to markets and supermarkets in the city (Figure 1).

The identification of a counterfactual, representing the situation without PGS certification, is important for attributing outcomes to standard adoption. We thus rely on a multistage sampling strategy: First, all 119 farmers from the PGS-Vietnam scheme growing
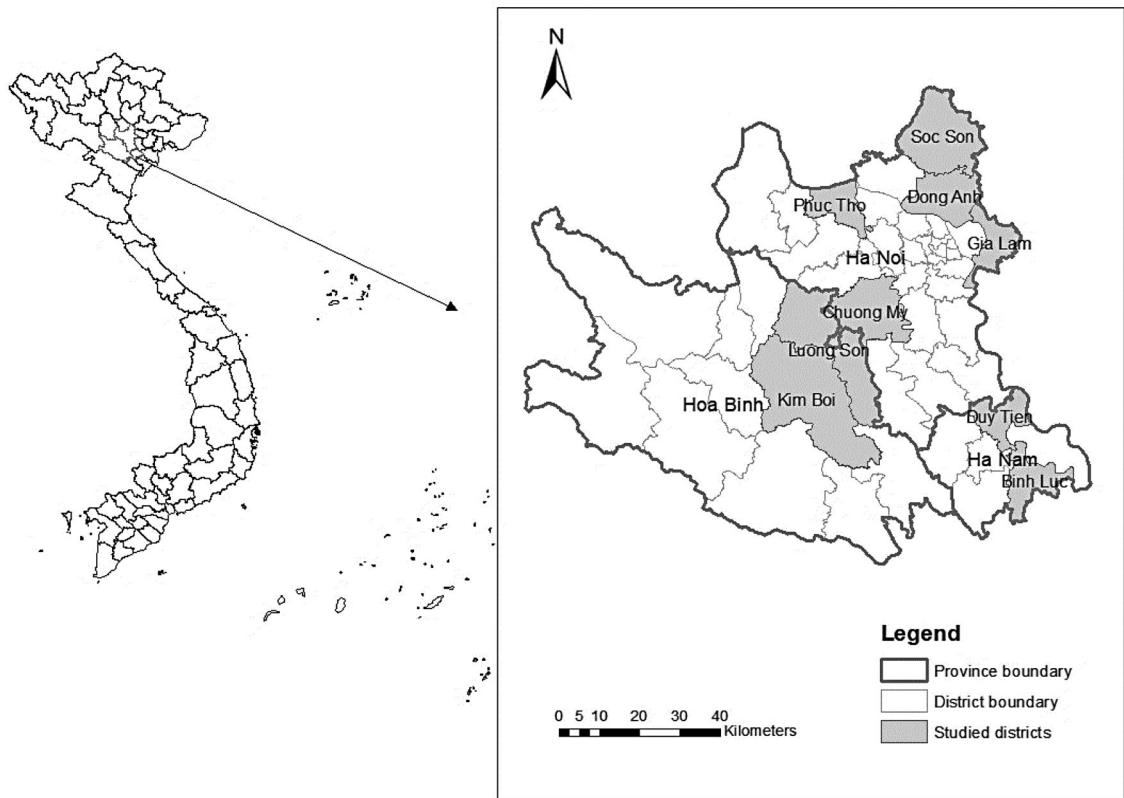
Figure 1. Map of the study area in Vietnam.

white cabbage were included in the sample for the intervention group. This represents 80% of the 149 PGS-certified farms around Hanoi. For the comparison group, power calculations on profitability, a key outcome variable, suggested a minimum group size of 279 farms (alpha = 0.05, power = 0.8, intervention group = 120 farms). We added 5% in case that some observations had to be discarded in the analysis. The power calculations were based on a one-sided two-means comparison and a fixed intervention group size of 120. Without similar studies in Vietnam or neighbouring countries, averages and standard deviations for organic and conventional cabbage profitability were based on expert opinion and compared to the global literature. The following approach was subsequently used to select farmers into the comparison group: for each of the 9 vegetable-producing districts, major white cabbage production villages (defined as villages with at least 50 farmers growing cabbage in a conventional way) were identified with the help of local extension officers. Next, twenty farmers were randomly selected from a list of all

cabbage farmers per village; and 10–14 of these were selected by village head for interviews based on their availability. We continued to add selected villages until we had completed 301 interviews. The total sample included 420 cabbage farmers spread over 26 villages in 9 districts, of which 119 farmers were PGS-certified.

The structured questionnaire included approximately 200 questions and was tested with farmers in a conventional and a PGS location before the start of the data collection. It was approved by the ethics committee of the Research Institute of Organic Agriculture (FiBL; Approval Nr. FSS-2023-07). The survey was implemented in the Kobo Toolbox, with answers recorded on tablets and exported into a customized MS Excel database. The respondent was identified as the farm manager responsible for vegetable production. Interviews took about 90 min on average. At the start of the interview, enumerators explained the purpose and content of the survey and ensured that data would be kept confidential. All respondents were asked for their verbal consent to participate.

Table 1. Summary statistics of the outcome variables used in the analysis.

Description of outcome variables	Short name of outcomes	Farms without PGS certification (n = 301)		Farms with PGS certification (n = 119)	
		Mean	SD	Mean	SD
Gross margin (USD/ha)	*gm*	6,036.25	4,192.27	11,812.57	7,278.75
Revenues (Million USD/ha)[a]		8,135.04	4,138.51	13,589.91	7,654.42
Variable input costs (USD/ha)[a]		2,082.69	1,209.01	1,765.34	1,451.44
Return to labour (USD/personday)	*returnlab*	22.27	20.20	15.60	12.69
Own labour (persondays/ha)[a]		430.54	417.66	1,179.23	953.85
Crop yield (t/ha)	*yield*	34.72	14.00	21.60	11.43
Agroecology performance (0–1 score)	*tape5*	0.31	0.08	0.47	0.07
Awareness of AE practices (# 1–11)	*aware*	6.95	2.91	9.06	1.34
Use of AE practices (# of 1–11)	*use*	4.02	2.34	6.90	1.57
Confidence of using AE practices (0–1 score)	*conf*	0.46	0.24	0.62	0.13
Knowledge (# of key words mentioned, 1–9)	*orgknow*	2.05	2.02	4.12	2.42
Choice of marketing channel (1 = yes)	*choice*	0.60	0.49	0.83	0.38

Note: AE = agroecological; gross margin, return to labour and yield are crop-specific (white cabbage), other variables are farm-specific.
[a]These variables were used in calculations but were not included in the analysis as outcome variables.

Data quality was checked daily by a designated data quality control officer. The anonymized data are available from the authors upon request.

The data included information about farm and farmer characteristics, certification, and production, including output quantities and values as well as and input use for white cabbage. Agroecological outcomes were measured using part of the Tool for Agroecology Performance Evaluation (TAPE; FAO, 2019), which was deemed relevant to Vietnamese farming context. Step 1 of TAPE operationalizes the ten elements of agroecology through a multi-criteria scoring approach. The ten elements of agroecology characterize the sustainability of agricultural production systems, highlighting their level of transition to agroecology. The TAPE diagnostic can be used to identify the strengths and weaknesses of these systems. For this purpose, each element is divided into multiple sub-elements, which are measured on a scale from zero to four.[2] The diversity score for instance contains ratings for the diversity of trees, livestock, trees, and other economic activities (FAO, 2019). All data were analysed using the statistical software Stata17. Monetary values were converted from Vietnamese Dong (VND) to US dollars (USD), while local area units were converted from sao (360 m^2) to hectares.[3] Extreme outliers in the cost, yield and price variables were winsorized at the 1st and 99th percentiles, with replacement.

2.3. Outcome and control variables

We consulted with experts from Vietnam, Switzerland, and FAO to define a broad set of sustainability outcomes for this study, as shown in Table 1. Variable choice is informed by the 13 principles of agroecology (Wezel et al., 2020) as well as FAO's ten elements of agroecology (FAO, 2018a). It follows an impact pathway logic leading from awareness to adoption, and further to improved farm performance.

2.3.1. Economic outcomes
White cabbage productivity and profitability were used as economic performance measures. Data were collected on farmers' expenditures on all inputs (organic and mineral fertilisers, synthetic and bio-based pesticides, growth regulators, hired labour, planting materials, etc.). The main variables of interest are white cabbage yield (in t/ha), gross margin (including hired labour cost) and return to labour, as the ratio between gross margin and all labour input. This is important to capture also differences in own labour, which, other than hired labour, is not valued at market prices.

2.3.2. Environmental outcomes
Eleven agroecological practices relevant to vegetable production in northern Vietnam were selected by the project team with questions about awareness, uptake and confidence (Figure 2A). As shown in Table 2, outcomes include the number of agroecological practices the respondent is aware of, the number of these practices used, and the respondent's confidence in these practices (0–1 score). While *aware* and *use* are count variables, *conf* is measured as a mean score on a 3-point scale, from no confidence (0) to high confidence (2) for each practice, with individual ratings averaged over all eleven practices and then standardized from zero to unity (0-1).
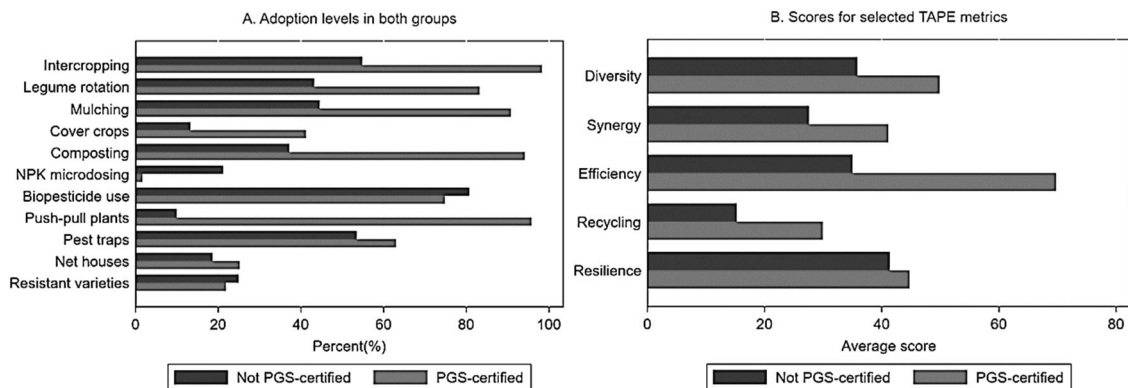
Figure 2. Descriptive group comparison for the individual components of aggregate environmental outcome variables 'use' (graph A) and 'tape5' (graph B).

FAO's TAPE indicator set was used to quantify the ecological performance of farms (FAO, 2019). We selected five out of ten metrics from TAPE Step 1: diversity, synergy, efficiency, recycling, and resilience (Figure 2B). The other five metrics (culture and food traditions, knowledge co-creation, human and social values, circular and solidarity economy, and responsible governance) are also important, but less relevant to measure environmental sustainability. We then created the aggregate outcome variable *tape5* and transformed it into a score ranging from zero to unity. This represents the average of five sustainability metrics.

2.3.3. Social outcomes
The two variables used in this category are both considered proxies for social empowerment. The first variable captures farmers' holistic understanding of organic farming. It is measured as the number of keywords mentioned when asking respondents to explain in their own words the meaning of organic farming. Eleven keywords were defined and could be marked.[4] The second variable captures choice vs. no choice of sales channel, asking farmers if they can only sell to one outlet or if they have a choice of multiple outlets. These two social outcomes indicate the strength of rural communities and the local economy.

Table 2. Description of agroecological practices used as outcomes.

Practice	Description
B1 Intercropping	Different crops are grown together on one plot, e.g. lettuce together with cabbage, cauliflowers, beans in a mosaic pattern or vegetables grown in-between early-stage fruit crops
B2 Rotation with legumes	Commercial legume crops, such as long bean, soybean, green beans, are rotated with vegetable crops according to cropping seasons.
B3 Mulching	Dry crop residues, such as rice, maize or legume husks) are used to cover the soil beds when growing vegetables.
B4 Cover crops	Fallow plots are covered by wild legumes or grasses, which farmers consider improves soil quality before the subsequent cropping season
B5 Microbial composting	Molasses and microorganism products are used for improving the composting process of manure and crop residues.
B6 Mineral fertilizer microdosing	Application of small quantities of mineral fertiliser at planting time or as top dressing after planting
B7 Biopesticides or natural predator release	Application of biological pesticides, such as *bacillus thuringiensis* and *pyrethroid,* which are purchased in the markets or of self-produced biological pesticides. Release of biological control products sold in the market, which contain predators, such as spiders or ladybugs.
B8 Push or pull crops (trap/ barrier crops)	Use of plants on the edge of the plot to either repel or trap insects, such as chrysanthemum, pot marigold, or garlic.
B9 Pest pheromone traps	Use of yellow or green sticky traps or molasses fly trap.
B10. Net houses for pest protection	Vegetable tunnels or greenhouses that are fully covered by plastics or tight nets to protect vegetables from insects and other animals.
B11. Pest resistant varieties	Use of pest resistant vegetable varieties, such as hybrid tomato F1 VNS390 or Japanese cucumber Gaga F1.

Table 3. Summary statistics of control variables in the analysis.

Description of covariates	Short name of covariates	Farms without PGS certification (n = 301)		Farms with PGS certification (n = 199)	
		Mean	SD	Mean	SD
Sex (1 = female)	sex	0.74	0.44	0.85	0.36
Age of farm manager (years)	age	58.02	9.03	55.70	8.68
Education of farm manager (years)	edu	7.47	2.20	7.83	2.03
Vegetable growing experience of farm manager (years)	vegexp	27.70	12.49	14.51	8.17
Household size (# of persons sharing a meal)	hhsize	4.30	2.07	4.72	1.83
Non-farm income (1 = very low, 2 = low, 3 = med, 4 = high)	nfinc	2.87	1.16	3.18	1.01
Loan (1 = yes)	loan	0.13	0.34	0.13	0.33
Farm size (ha)	fsize	0.42	0.21	0.42	0.23
Proportion of rented farmland (%)	rentp	0.38	0.17	0.26	0.17
Size of vegetable area (ha)	vegsize	0.16	0.11	0.15	0.19
Proportion of vegetable area with title (%)	titlep	0.80	0.42	0.83	0.69
Extension visit in the last five years (1 = yes)	ext	0.06	0.24	0.10	0.30
Interaction with neighbours (1 = no, 2 = some, 3 = regular)	neighb	2.40	0.74	2.73	0.62
Membership in farmer organization (1 = yes)	forg	0.49	0.50	0.56	0.49

2.3.4. Controls

Fourteen covariates conceptually related to PGS participation were selected based on previous studies (e.g. Gautam et al., 2017; Meemken & Qaim, 2018a) and the expertise of the researchers: sex, age, education and vegetable growing experience of the farm manager, household size (in terms of persons sharing a meal), non-farm income, loan from bank, farm size, the proportion of rented farmland, size of vegetable area, proportion of vegetable area with title, training attended in the last five years, extension visit in the last five years (1 = yes), frequency of contact with neighbours, membership in the national farmer organization (Table 3).

2.4. Econometric estimation

Rather than determining the average effect at the level of the entire population, our interest is to understand the effectiveness of certification for the farmers who took part in and complied with the PGS standard. Therefore, we estimate the average treatment effect on the treated (ATT), which is the average impact on those subjects who ultimately participate in the intervention (Austin, 2011; Imbens & Wooldridge, 2009). Contrary to a weighted mean comparison for ATT estimation, the advantage of a weighted regression adjustment estimator is that it provides correct standard errors for the ATT (Jann, 2021). This is a function of the doubly robust properties of this estimator (Kurz, 2022; Zhao & Percival, 2016). The weighted regression adjustment implies that weights are determined in the first step for

subsequent inclusion in an outcome model in the second step. Both steps include the use of the same covariates (Jann, 2021).

Weighting has advantages over the more widely used propensity score matching since observations without an exact or approximate match are not dropped from the analysis (Hirano et al., 2003). This allows for maintaining the full comparison group (Meemken & Qaim, 2018a). We used entropy weights and inverse probability weights as alternative approaches. Both rely on the same set of covariates, as shown in Table 3, and address systematic differences in the distribution of these observed covariates between the intervention group of certified farms and the comparison group of non-certified farms. The covariates in Table 3 are assumed to simultaneously affect the certification decision and the outcome variables and are used to compute the weights. When systematic differences are controlled for through weighting, the two groups become similar to the extent that remaining differences in the outcome variables can be interpreted as 'treatment effects' (Meemken & Qaim, 2018a).

Entropy balancing, as proposed by Hainmueller (2012), generates weights that exactly balance all covariates according to their mean and higher moments if convergence is achieved. The method is implemented in Stata through the *ebalance* routine (Hainmueller & Xu, 2013). Weights are assigned to each farm in the comparison group, but not to the farms in the intervention group, resulting in the estimation of the ATT as $E(Y_1 - Y_0|U_i = 1) = E(Y_1|T_i = 1) - E(Y_0|U_i = 1)$, where Y_1 is the outcome with certification, Y_0 is the outcome

without certification and T_i refers to farmers' participation in the PGS scheme. Entropy balancing generally achieves higher estimation accuracy than inverse probability weighting (Harvey et al. 2017). When computing the entropy weights in our analysis, convergence was achieved by balancing all covariates for both mean and variance. This produced almost identical statistics for the intervention and comparison groups, while inverse probability weights provided only approximate balance between groups (appendix, Table A). As inverse probability weights are commonly used in ex-post impact evaluation and are slightly less prone (as compared to entropy weights) to contain extreme weighting values for individual observations, they were used to verify the results. Inverse probability weights are generated by predicting the propensity score (probability of treatment) from a logit regression (Austin, 2011), including the covariates from Table 2 as predictors. Weights were calculated as $1/(1 -$ propensity score) for the untreated group, while the treated group was not weighted when estimating the ATT. To obtain correct standard errors, the computed weights were entered into the regression adjustment model, which is run for each set of weights, using the same covariates as used in the first step of the data processing (appendix, Table B). Regression adjustment was performed using the *teffects ra* command in Stata according to the distribution of the outcome variables. Linear outcome models were specified for the variables *gm*, *labreturn* and *yield*. For the count variables *aware*, *use* and *orgknow*, Poisson models were applied. Fractional outcome models were used for the variables *tape5* and *conf*, which are bounded between zero and unity, while a binary outcome model was used for the outcome variable *choice*. We assumed that observations within a district are not completely independent from one another and therefore clustered the standard errors by district.

2.5. Robustness checks

As a robustness check, we assessed the extent to which our results are sensitive to omitted variable bias as weighted regression adjustment does not control for the risk of this bias. This may be an issue in our study because it is not possible to be aware of and collect data on all possible confounders. For instance, information on certain personal characteristics of farmers (e.g. attitudes and beliefs), that might influence the outcomes of interest, is not available and might influence the results. To understand

the potential effect of omitted variables on our results, we applied the regression sensitivity approach proposed by Diegert et al. (2022), implemented in Stata as *regsensitivity*. Contrary to Oster bounds for selection on unobservables (Oster, 2019), it allows for the included control variables to be correlated with unobserved control variables, comparing the magnitude of selection on observables with the magnitude of selection on unobservables. A so-called breakdown point (*rxbar*) defines the threshold for the results to be overturned by omitted variable bias (Diegert et al., 2022). The higher the breakdown point, the less likely it is for the results to be heavily influenced by unobserved factors. Using the same covariates as in the main treatment effect estimations, we applied this robustness check to all outcome variables. Breakdown points are calculated for five levels of correlation (*cbar*) between observed and unobserved variables (0.00, 0.25, 0.50, 0.75, 1.00).

We also verified the results to understand if there are underlying structural biases. To reflect the importance of different income sources, we used subsamples of the variable measuring on-farm income. The robustness of the results was checked when excluding farms with low on-farm income (1) as well as those with high on-farm income (4).

3. Results

Significant positive results were obtained for most outcome variables, with effect sizes ranging from 23% to over 100% (Table 4). This demonstrates that organic PGS has a relevant role in making vegetable production more viable and more agroecological. However, some key challenges also became evident as there are large and negative effect sizes for crop yield and return to labour, albeit only significant for yield. The estimated treatment effects are very similar in terms of direction and significance for both approaches, which gives confidence to the results. Nevertheless, some of the effect sizes differ, e.g. for crop yield or knowledge of organic farming. This can be explained by the fact that entropy weights can be more extreme and provide exact balance among groups, while the balance is more approximate in the case of inverse probability weights, where observations are less likely to receive extremely small weights (Table 4).

PGS has a large and positive impact on the profitability of white cabbage production, as shown by the results for *gm*. However, the return to labour is not significant for both estimations as PGS uses

Table 4. Estimated treatment effects.

Outcomes	Unit	Entropy weighted regression adjustment					f	Inverse probability weighted regression adjustment				
		ATT	Sign	SE	POM	ATT (%)		ATT	Sign	SE	POM	ATT (%)
gm	USD/ha	6,377.96	***	(1,607.38)	5,434.61	117.4		5,671.48	***	(1,538.220)	6,141.08	92.4
returnlab	USD/day	−355.08	ns	(254.060)	1915.45	−18.54		−440.10	ns	(324.55)	2000.34	−21.9
yield	t/ha	−8.71	***	(2.563)	30.24	−28.8		−13.38	***	(2.541)	34.92	−38.3
tape5	0–1 score	0.14	***	(0.015)	0.33	40.6		0.15	***	(0.016)	0.32	46.2
aware	# practices	2.15	***	(0.247)	6.90	31.2		2.54	***	(0.225)	6.52	39.0
use	# practices	2.79	***	(0.493)	4.11	67.8		2.70	***	(0.517)	4.19	64.5
conf	0–1 score	0.15	***	(0.041)	0.47	32.6		0.19	***	(0.028)	0.43	45.0
orgknow	# keywords	1.59	***	(0.388)	2.52	63.1		2.10	***	(0.222)	2.02	103.7
choice	1 = yes	0.16	*	(0.090)	0.67	23.3		0.21	***	(0.076)	0.62	33.2

Note: See Table 2 for full names of outcome variables.
ATT, average treatment effect on the treated; ATT (%) is calculated as ATT/POM. POM, potential outcome mean; SE, standard error; Sign., significance.
***Significance level: 0.01.
**Significance level: 0.05.
*Significance level: 0.1.

substantially more labour. Crop yield is negatively affected by PGS participation with the effect size ranging from −29% to −38%.

Environmental performance of PGS participation is measured through five elements from FAO's TAPE score as well as through awareness and adoption of agroecological practices. This is complemented by a confidence score for these practices. Across all outcome variables (tape5, aware, use and conf) we determined a positive and significant impact of PGS participation, ranging from 41% to 46% for the TAPE score and from 33% to 45% for the confidence score.

Knowledge of organic farming, measured in terms of mentioned keywords and choice of market channel, is used as a proxy for social empowerment of farmers. The effect sizes are positive and significant, ranging from 63% to 104% for knowledge and from 23% to 33% for choice of market channel.

As robustness checks, subsampling and omitted variable tests were carried out. Initially, we performed the econometric estimations with two subsets of the data. Farms with low non-farm income were excluded, so that the remaining sample contained 108 PGS-certified farms and 243 non-PGS farms. This resulted in very similar effect sizes and, apart from the variable choice, there was no chance in significance levels. When farms with the highest on-farm incomes were excluded, we observed the same results, having 101 PGS-certified farms and 248 farms in the comparison group. In addition to subsets of the data, sensitivity tests were carried out to assess the risk of omitted
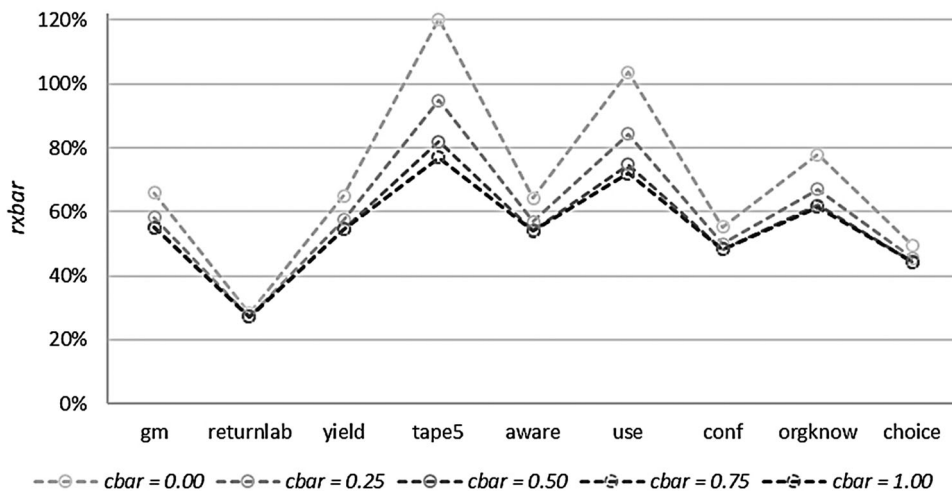
Figure 3. Sensitivity analysis with breakdown points (rxbar) for different correlation levels (cbar).

variable bias for all outcome variables in the analysis. The breakdown points (*rxbar*) shown in Figure 3 indicate at what point the direction of our estimated treatment effects will be overturned (also see appendix Table A). The percentages for each outcome variable give us an indication of the extent of selection on unobservables, as compared to selection on observables, which would result in this reversal. A range of breakdown points were computed for different levels of correlation (cbar) between unobserved and observed variables, with 1 indicating full correlation and 0 indicating no correlation. Likely some slight to moderate correlation can be expected. No clear cut-off exists for ruling out the risk of omitted variable bias through the regression sensitivity approach. The method rather helps to quantify the risk, with low breakdown points indicating a high likelihood of selection on unobservables occurring alongside selection on observables. As we have access to a comprehensive set of co-variates, we do not expect that more than 50% of the selected bias corrected by observables would be caused by unobserved factors. We consider this a reasonable assumption and conclude that estimations for two variables (*returnlab* and *choice*) might be prone to bias from omitted variables, with the return to labour analysis more seriously affected than the analysis of the sales channel choice. It appears that our impact model for return to labour might be missing some variables that were not captured in the survey. Yet, judging from the descriptive data for labour and gross margin, the general trend of the result for return to labour seems plausible (appendix, Table B).

4. Discussion

As the first impact evaluation of organic PGS certification, our analysis confirms the original premise that organic PGS certification delivers economic, environmental, and social benefits. The impact of PGS is positive across all indicators of agroecology uptake and agroecological performance but mixed for the economic indicators. Gross margins were positively affected by PGS participation, but high labour requirements and a substantial yield penalty reduced the economic viability and scalability of current PGS-certified vegetable production in northern Vietnam. Robustness of the findings is confirmed by using two weighting approaches, checking sensitivity to omitted variables, and examining the underlying structural biases. However, the focus on one

season and one crop for the quantification of economic outcomes may somewhat affect the validity of the results. While white cabbage is the most important crop for vegetable farmers in the study area, the overall economic performance of farms depends on a larger range of crops.

For profitability, our results are in line with evidence from meta-analyses (Crowder & Reganold, 2015; Seufert & Ramankutty, 2017), but do not confirm the findings by Pham and Shively (2019). The latter study, examining upland vegetable production in northern Vietnam, showed that organic certification did not increase farm profitability. Their analysis focused on the third-party TCVN standard and used OLS regression or a relatively small sample covering multiple crops. Our evaluation is focused on the main commercial crop in the study area for the winter cropping season and relies on a more robust counterfactual analysis. The highly significant positive effect on white cabbage profits, which the analysis revealed, is driven by considerable price premiums for organic vegetables and by lower variable input costs as farmers use fewer chemical inputs.

Meemken and Qaim (2018b) pointed out that many studies on the profitability of organic agriculture in low- and middle-income countries do not account for the opportunity cost of family labour. Our study addresses this issue by including returns to labour as an outcome variable. Once own labour is accounted for, the positive effect on farm profitability disappears. The review by Crowder and Reganold (2015) found that organic agriculture uses 13% more labour on average. According to their analysis, when factoring in labour costs then the benefit–cost ratios of organic vegetables are lower than for all other crops but on par with those of conventionally produced vegetables. For vegetable production in Europe, labour requirements have been reported to be twice as high on organic vegetable farms as on conventional farms (Jansen, 2000). We could not find comparative studies on labour use on vegetable farms in Asia, but our descriptive data suggest that organic PGS farms use 2.5 times more labour than conventional farms. When considering labour tasks, we find that the labour input on PGS-certified farms is higher for land preparation, harvesting and sorting, and especially for compost application and weeding. Labour shortages will challenge the economic viability of organic PGS vegetable production, but organic PGS also creates jobs in areas with fewer employment opportunities. As many farmers

are close to retirement age and labour markets are highly competitive in peri-urban contexts such as Hanoi where youth have opportunities outside agriculture, labour-saving technologies are needed for PGS organic development. Expert opinions suggest that improved seed for organic farming conditions, exact transplanting of seedlings and less soil disturbance through reduced tillage can be promising strategies for reducing the number of person-days required in land preparation and weeding. Furthermore, collective crop planning and growing schemes for PGS vegetable cultivation among farmer group members have the potential to facilitate shared input use and plot size optimization, thus reducing labour requirements.

For vegetable yields, reviews suggest a gap ranging between 13% and 33% (see Table 2 in Meemken & Qaim, 2018b). Based on farm data rather than experimental data analysis, Kniss et al. (2016) reported yield gaps of 50% and more for vegetables in the US. For Asia, we found no research on vegetable yield gaps, but the meta-analysis of De Ponti et al. (2012) suggests 11% across major crops. For our case, yield reduction due to participation in organic PGS is at the upper end of the range that meta-analyses suggest. To the knowledge of the authors, this can be mainly explained by the absence of synthetic inputs in organic PGS-certified production systems, a lack of availability of organic vegetable seed and suboptimal soil management. Reduced tillage and better organic amendments can help improve soil structure, organic input effectiveness, and yields in the PGS vegetable systems. From an economic perspective, high price premiums compensate for lower yields. However, not many consumers can, on a regular basis, afford organic produce that is twice or even three times as expensive as conventionally produced vegetables. Therefore, a reduced yield gap and lower labour requirements in PGS-certified production systems may also translate into more affordable organic vegetables, with price premiums below a 100% markup. The study by Hai et al. (2013) suggested that consumers in Hanoi are willing to pay 70% more for organic vegetables. Besides productivity increases, direct marketing can also help in lowering the cost of organic vegetables for consumers and improve farmers' income (Wang et al., 2014). Our results show that PGS participation leads to more choice in terms of marketing channels, however, direct marketing to consumers is done by very few PGS farmers. Digital tools and collective arrangements can help to develop this option into a more widely used sales outlet. Navarette et al. (2015) point out that crop management, labour organization, and marketing are crucial issues for the sustainability of organic vegetable production. While the findings are based on data from France, the same challenges seem to also apply to our case study in Vietnam.

Despite lower productivity, organic farming has been shown to provide a range of environmental benefits, such as increased biodiversity (Gong et al., 2022) and resilience. To determine impacts on environmental sustainability, a range of outcomes related to agroecological growing practices were evaluated. In addition, five elements of the TAPE tool were used in the analysis to assess the effects of organic PGS certification on agroecology performance. Substantial positive changes in awareness, uptake and confidence regarding practices, as well as an increase in the TAPE score, can be attributed to PGS participation. This demonstrates that organic vegetable production based on credible participatory certification can generate clear benefits in terms of environmental sustainability. The scoring results from the TAPE assessment are confirmed by the data on the uptake of agroecological practices. While the application of TAPE is generally straightforward, the wording for scores needed to be contextualized and not all scoring categories are always applicable to a specific case. In addition, compound statements can complicate the exact measurement of concepts.[5] While in some instances clarity of scores can be improved, the overall application of a selection of TAPE elements has proved to be a valuable addition to the set of outcome indicators, allowing for a better understanding of the impacts of PGS certification on the sustainability of vegetable production in our study area. A more complete application of TAPE, including metrics, such as knowledge co-creation, human and social values or circular and solidarity economy, would allow researchers to better capture social outcomes of certification. Empowerment and revitalization of rural communities may thus be considered in future impact studies on certification standards.

5. Conclusion

Organic PGS standards have been around for more than two decades and are increasingly used as a certification tool in middle-income countries, such as Vietnam, where there is a rising consumer demand

for safe food. This is the first quantitative evaluation of the impact of organic PGS certification in Vietnam. The article is based on a counterfactual analysis, state-of-the-art robustness checks and a broad set of sustainability outcomes. The study shows that the production of PGS-certified organic vegetables in Vietnam leads to higher profits and better agroecological performance, but also faces challenges, such as lower crop yield and higher labour requirements, which are effects consistent with studies for other countries. While these trade-offs are well established, they are not a given and it remains important to quantify the extent of positive and negative effects of PGS certification.

To attract more farmers, especially youth, and realize the agroecological potential of organic PGS production systems, the challenges can be addressed through improved advisory services for PGS farmers. Our study identifies several areas for improvements in organic PGS vegetable production, which can be integrated into capacity development interventions. Suggested improvements involve transplanting of healthy seedlings rather than direct seeding as well as applying reduced tillage for enhanced soil structure, nutrient holding capacity and weed suppression. Collective crop planning and management are also important to enhance labour efficiency, farming structure and overall farming landscape ecosystem services. Government and other stakeholders' investments and efforts to promote these practices appear well justified, considering the positive environmental externalities and food safety advantages obtained in PGS-certified vegetable production. To reduce uncertainty, labour-saving techniques and improved soil fertility management need to first be piloted among selected organic vegetable farmers. Future research should then systematically evaluate benefits and risks before more widely scaling recommended practices.

Notes

1. Valid instruments need to be correlated with the intervention (such as standard participation), but must not be correlated with the outcomes, except through their indirect effect on the intervention (White & Raitzer, 2017).
2. The TAPE methodology and examples of the tool's application are available on the TAPE website of FAO: https://www.fao.org/agroecology/tools-tape/en/
3. At the time of the survey, the prevailing exchange rate was 23,500 VND = 1 USD.

4. Such as: no synthetic pesticides, no chemical fertilisers, natural nutrient cycles, natural inputs, beneficial organisms, local seeds, crop diversity, balanced agroecosystems, premium prices, specialty markets, other.
5. E.g. the description for the highest score of 'Crop diversity' under the 'Diversity' element: 'more than 3 crops and varieties adapted to local conditions. Spatially diversified farm by multi-, poly- or intercropping'.

ORCID

Pham Van Hoi http://orcid.org/0000-0003-3290-3012

References

Altieri, M. A. (1995). *Agroecology: The science of sustainable agriculture* (2nd ed.). Boulder, CO: Westview Press.

Amekawa, Y. (2013). Can a public GAP approach ensure safety and fairness? A comparative study of Q-GAP in Thailand. *The Journal of Peasant Studies*, *40*, 189–217.

Austin, P.C. (2011). An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research*, *46*(3), pp. 399–424.

Blockeel, J., Schader, C., Heidenreich, A., Grovermann, C., Kadzere, I., Egyir, I.S., Muriuki, A., Bandanaa, J., Tanga, C.M., Clottey, J., & Ndungu, J. (2023). Do organic farming initiatives in Sub-saharan Africa improve the sustainability of smallholder farmers? Evidence from five case studies in Ghana and Kenya. *Journal of Rural Studies*, *98*, 34–58.

Bolwig, S., Gibbon, P., & Jones, S. (2009). The economics of smallholder organic contract farming in tropical Africa. *World Development*, *37*, 1094–1104.

CFS. (2021). *Report of the 49th session: Making a difference in food security and nutrition*. Committee on World Food Security.

Crowder, D.W., & Reganold, J.P. (2015). Financial competitiveness of organic agriculture on a global scale. *Proceedings of the National Academy of Sciences*, *112*, 7611–7616.

De Ponti, T., Rijk, B., van Ittersum, M. K. (2012). The crop yield gap between organic and conventional agriculture. *Agricultural Systems*, *108*, 1–9.

Diegert, P., Masten, M. A., & Poirier, A. (2022). Assessing omitted variable bias when the controls are endogenous. arXiv: 2206.02303.

FAO. (2018a). *10 elements of agroecology*. UN Food and Agriculture Organization.

FAO. (2018b). *Asia and the Pacific regional overview of food security and nutrition 2018 – accelerating progress towards the SDGs*. UN Food and Agriculture Organization.

FAO. (2019). *TAPE tool for agroecology performance evaluation – process of development and guidelines for application. Test version*. UN Food and Agriculture Organization.

Gautam, S., Schreinemachers, P., Uddin, M. N., & Srinivasan, R. (2017). Impact of training vegetable farmers in Bangladesh in integrated pest management (IPM). *Crop Protection*, *102*, 161–169.

Gliessman, S. (2018). Defining agroecology. *Agroecology and Sustainable Food Systems*, *42*, 599–600.

Gong, S., Hodgson, J.A., Tscharntke, T., Liu, Y., van der Werf, W., Batáry, P., Knops, J.M., & Zou, Y. (2022). Biodiversity and yield trade-offs for organic farming. *Ecology Letters*, *25*, 1699–1710.

Grovermann, C., Quiédeville, S., Muller, A., Leiber, F., Stolze, M., & Moakes, S. (2021). Does organic certification make economic sense for dairy farmers in Europe?–a latent class counterfactual analysis. *Agricultural Economics*, *52*, 1001–1012.

Grovermann, C., Schreinemachers, P., & Berger, T. (2013). Quantifying pesticide overuse from farmer and societal points of view: An application to Thailand. *Crop Protection*, *53*, 161–168.

Ha, T.M., Shakur, S., & Do, K.H. (2019). Consumer concern about food safety in Hanoi, Vietnam. *Food Control*, *98*, 238–244.

Hai, N.M., Moritaka, M., & Fukuda, S. (2013). Willingness to Pay for organic vegetables in Vietnam: An empirical analysis in Hanoi capital, journal of faculty of agriculture. *Kyushu University*, *58*, 449–458.

Hainmueller, J.(2012). Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis 20*, 25–46.

Hainmueller, J., & Xu, Y. (2013). Ebalance: A Stata package for entropy balancing. *Journal of Statistical Software*, *54*, 1–18.

Harvey, R.A., Hayden, J.D., Kamble, P.S., Bouchard, J.R., & Huang, J.C. (2017). A comparison of entropy balance and probability weighting methods to generalize observational cohorts to a population: A simulation and empirical example. *Pharmacoepidemiology and Drug Safety*, *26*, 368–377.

Hirano, K., Imbens, G.W., & Ridder, G. (2003). Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica*, *71*, 1161–1189.

HLPE. (2019). *Agroecological and other innovative approaches for sustainable agriculture and food systems that enhance food security and nutrition*. High Level Panel of Experts on Food Security and Nutrition of the Committee on World Food Security.

Home, R., Bouagnimbeck, H., Ugas, R., Arbenz, M., & Stolze, M. (2017). Participatory guarantee systems: Organic certification to empower farmers and strengthen communities. *Agroecology and Sustainable Food Systems*, *41*, 526–545.

Ibanez, M., & Blackman, A., 2016. Is eco-certification a win–win for developing country agriculture? Organic coffee certification in Colombia. *World Development*, *82*, 14–27.

IFOAM. (2019). *PGS guidelines. How to develop and manage participatory guarantee systems for organic agriculture*. IFOAM Organics International. Bonn.

Imbens, G.W., & Rubin, D.B., 2015. *Causal inference in statistics, social, and biomedical sciences*. Cambridge University Press.

Imbens, G.W., & Wooldridge, J.M. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature*, *47*, 5–86.

IPES-Food. 2016. *From uniformity to diversity: A paradigm shift from industrial agriculture to diversified agroecological systems. International panel of experts on sustainable food systems*. Louvain-la-Neuve.

Jann, B. (2021). Entropy balancing as an estimation command. Working Paper No. 39. University of Bern Social Sciences.

Jansen, K. (2000). Labour, livelihoods and the quality of life in organic agriculture in Europe. *Biological Agriculture & Horticulture*, *17*, 247–278.

Kerr, R.B., Madsen, S., Stüber, M., Liebert, J., Enloe, S., Borghino, N., Parros, P., Mutyambai, D.M., Prudhon, M., & Wezel, A. (2021). Can agroecology improve food security and nutrition? A review. *Global Food Security*, *29*, 100540.

Kniss, A. R., Savage, S. D., & Jabbour, R. (2016). Commercial crop yields reveal strengths and weaknesses for organic agriculture in the United States. *PLoS One*, *11*, e0161673.

Knook, J., Eory, V., Brander, M., & Moran, D., 2018. Evaluation of farmer participatory extension programmes. *The Journal of Agricultural Education and Extension 24*, 309–100325.

Kurz, C. F. (2022). Augmented inverse probability weighting and the double robustness property. *Medical Decision Making*, *42*, 156–167.

Loconto, A., & Hatanaka, M. (2018). Participatory guarantee systems: Alternative ways of defining, measuring, and assessing 'sustainability'. *Sociologia Ruralis*, *58*, 412–432.

Lopez-Avila, D., Husain, S., Bhatia, R., Nath, M., & Vinaygyam, R. (2017). Agricultural innovation: An evidence gap map. 3ie Evidence Gap Map Report 12. New Delhi, International Initiative for Impact Evaluation (3ie).

Meemken, E.M., & Qaim, M. (2018a). Can private food standards promote gender equality in the small farm sector? *Journal of Rural Studies*, *58*, 39–51.

Meemken, E. M., & Qaim, M. (2018b). Organic agriculture, food security, and the environment. *Annual Review of Resource Economics*, *10*, 39–63.

Mendoza, T.C. (2004). Evaluating the benefits of organic farming in rice agroecosystems in the Philippines, *Journal of Sustainable Agriculture 24*, 93–115.

Navarrete, M., Dupré, L., & Lamine, C., 2015. Crop management, labour organization, and marketing: Three key issues for improving sustainability in organic vegetable farming. *International Journal of Agricultural Sustainability*, *13*, 257–274.

Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics 37*, 187–204.

Pham, L., & Shively, G. (2019). Profitability of organic vegetable production in northwest Vietnam: Evidence from Tan Lac district, Hoa Binh Province. *Organic Agriculture*, *9*, 211–223.

Praneetvatakul, S., Schreinemachers, P., Pananurak, P., & Tipraqsa, P., 2013. Pesticides, external costs and policy options for Thai agriculture. *Environmental Science & Policy*, *27*, 103–113.

Schafer, J. L., & Kang, J. (2008). Average causal effects from non-randomized studies: A practical guide and simulated example. *Psychological Methods*, *13*, 279–313.

Schleifer, P., & Sun, Y. (2020). Reviewing the impact of sustainability certification on food security in developing countries. *Global Food Security*, *24*, 100337.

Schreinemachers, P., Afari-Sefa, V., Heng, C.H., Dung, P.T.M., Praneetvatakul, S., & Srinivasan, R., 2015. Safe and sustainable crop protection in Southeast Asia: Status, challenges and policy options. *Environmental Science & Policy*, *54*, 357–366.

Schreinemachers, P., Grovermann, C., Praneetvatakul, S., Heng, P., Nguyen, T.T.L., Buntong, B., Le, N.T., & Pinn, T. (2020). How much is too much? Quantifying pesticide overuse in vegetable production in Southeast Asia. *Journal of Cleaner Production*, *244*, 118738.

Schreinemachers, P., Schad, I., Tipraqsa, P., Williams, P.M., Neef, A., Riwthong, S., Sangchan, W., & Grovermann, C. (2012). Can public GAP standards reduce agricultural pesticide use? The case of fruit and vegetable farming in northern Thailand. *Agriculture and Human Values*, 29, 519–118529.

Seufert, V., & Ramankutty, N. (2017). Many shades of gray—the context-dependent performance of organic agriculture. *Science Advances*, 3(3), e1602638.

Ssebunya, B., Morawetz, U., Schader, C., Stolze, M., Schmid, E. (2019). Group membership and certification effects on incomes of coffee farmers in Uganda. *European Review of Agricultural Economics 46*, 109–132.

Tran, D., & Goto, D. (2019). Impacts of sustainability certification on farm income: Evidence from small-scale specialty green tea farmers in Vietnam. *Food Policy 83*, 70–82.

UNSG. (2020). *Report of the secretary-general: Progress towards the sustainable development goals*. New York: United Nations Economic and Social Council.

USDA. (2023). Vietnam issues national action plan on food systems transformation toward transparency responsibility and sustainability by 2030. United States Department of Agriculture – Foreign Agricultural Service Report VM 2023-0017.

van der Ploeg, J.D., Barjolle, D., Bruil, J., Brunori, G., Costa Madureira, L.M., Dessein, J., … , Wezel, A. (2019). The economic potential of agroecology: Empirical evidence from Europe. *Journal of Rural Studies 71*, 46–61.

von Braun, J., & Hendriks, S. (2023). Full-cost accounting and redefining the cost of food: Implications for agricultural economics research. *Agricultural Economics 54*, 451–454.

Wang, H., Moustier, P., & Loc, N.T.T. (2014). Economic impact of direct marketing and contracts: The case of safe vegetable chains in northern Vietnam. *Food Policy*, 47, 13–23.

Wanwimolruk, S., Phopin, K., Boonpangrak, S., & Prachayasittikul, V. (2016). Food safety in Thailand 4: Comparison of pesticide residues found in three commonly consumed vegetables purchased from local markets and supermarkets in Thailand. *PeerJ*, 4, e2432.

Wezel, A., Herren, B.G., Kerr, R.B., Barrios, E., Gonçalves, A.L.R., & Sinclair, F. (2020). Agroecological principles and elements and their implications for transitioning to sustainable food systems: A review. *Agronomy for Sustainable Development*, 40, 1–e2413.

White, H., & Raitzer, D.A. (2017). *Impact evaluation of development interventions: A practical guide*. Asian Development Bank.

Willer, H., Schlatter, B., & Trávníček, J. (Eds.). (2023). *The world of organic agriculture. Statistics and emerging trends 2023*. Research Institute of Organic Agriculture FiBL and IFOAM – Organics International. Frick and Bonn.

Zhao, Q. & Percival, D. (2016). Entropy balancing is doubly robust. *Journal of Causal Inference 5*(1).

Appendix

Table A. Balancing of covariates for both weighting approaches.

Covariates Name	Unit	Entropy weights				C	Inverse Probability Weights			
		cFarms without PGS certification (n = 301)		Farms with PGS certification (n = 199)			Farms without PGS certification (n = 301)		Farms with PGS certification (n = 199)	
		Mean	SD	Mean	SD		Mean	SD	Mean	SD
sex	(1 = female)	0.863	0.119	0.863	0.119		0.863	0.119	0.889	0.099
age	(years)	56.140	77.270	56.140	77.270		56.137	77.268	56.158	116.904
edu	(years)	7.600	3.626	7.600	3.626		7.600	3.626	7.655	4.562
vegexp	(years)	15.010	74.690	15.010	74.690		15.011	74.691	13.080	105.978
hhsize	(persons)	4.811	3.283	4.811	3.283		4.811	3.283	5.145	4.287
nfinc	(1-4)	3.358	0.594	3.358	0.594		3.358	0.594	3.467	0.497
loan	(1 = yes)	0.084	0.078	0.084	0.078		0.084	0.078	0.138	0.120
fsize	(ha)	0.422	0.053	0.422	0.053		0.422	0.053	0.392	0.046
rentp	(%)	0.272	0.028	0.272	0.028		0.272	0.028	0.274	0.028
vegsize	(ha)	0.155	0.037	0.155	0.037		0.154	0.037	0.135	0.010
titlep	(%)	0.847	0.555	0.847	0.555		0.847	0.555	0.786	0.139
ext	(1 = yes)	0.105	0.095	0.105	0.095		0.105	0.095	0.070	0.065
neighb	(1-3)	2.716	0.397	2.716	0.397		2.716	0.397	2.621	0.339
forg	(1 = yes)	0.558	0.249	0.558	0.248		0.558	0.249	0.556	0.248

Note: Full variable names are provided in Table 3.

Table B. Exact results of sensitivity analysis with breakdown points for different correlation levels.

Outcome	rxbar (Breakdown)				
	cbar = 0.00	cbar = 0.25	cbar = 0.50	cbar = 0.75	cbar = 1.00
gm	65.9%	58.1%	55.1%	55.0%	55.0%
returnlab	28.4%	27.3%	27.3%	27.3%	27.3%
yield	64.9%	57.4%	54.5%	54.5%	54.5%
tape5	120.3%	94.8%	82.0%	76.9%	76.9%
aware	64.2%	56.9%	54.1%	54.1%	54.0%
use	103.6%	84.4%	74.8%	71.9%	71.9%
conf	55.2%	49.9%	48.3%	48.3%	48.3%
orgknow	77.9%	67.0%	62.1%	61.5%	61.5%
choice	49.5%	45.4%	44.4%	44.4%	44.4%