

The effectiveness of agricultural certification in developing countries: a systematic review

Carlos Oya^a, Florian Schaefer^a, and Dafni Skolidou^b

Accepted version (pre-print).

Published in *World Development* Volume 112, December 2018, Pages 282-312

<https://doi.org/10.1016/j.worlddev.2018.08.001>

ABSTRACT

Certification systems (CS) set and monitor voluntary standards to make agricultural production sustainable in socio-economic terms and agricultural trade fairer for producers and workers. They try to achieve a wide range of socio-economic and environmental effects through bundles of interventions that include the process of standard setting and compliance, advocacy among consumers, capacity building for producers, building supply chains, price interventions, and the application of acceptable labour standards, overall to improve the wellbeing of farmers and agricultural workers.

This paper presents the results of a mixed-method systematic review that synthesized the literature on socio-economic effects of certification systems on agricultural producers and wage workers in low and middle income countries. The review followed the Campbell Collaboration guidelines for systematic reviews, and included studies published between 1990 and 2016 in different languages, with evidence on low and middle income countries. The review included a quantitative effectiveness question focused on a range of intermediate (e.g. prices, wages) and endpoint outcomes (e.g. household income). It also included a question on barriers, facilitators and contextual factors shaping effectiveness which drew on qualitative or mixed-method studies. Eligible certification systems were based on second- (industry-level) or third-party certifications, and excluded own-company standards. For the effectiveness review, quantitative impact evaluations must use experimental or non-experimental methods demonstrating control for selection bias. With these inclusion criteria, the review includes 43 studies used for analysing quantitative effects, and 136 qualitative studies for synthesizing barriers, enablers and other contextual factors. Most included studies report on initiatives in Latin America and sub-Saharan Africa and focus primarily on agricultural producers. The quality of the included studies is mixed, and several studies are weak on a number of methodological fronts, especially on statistical reporting.

Overall, there is limited and mixed evidence on the effects of CS on a range of intermediate and final socio-economic outcomes for agricultural producers and wage workers. There are positive effects on prices and income from the sale of produce is higher for certified farmers. However, workers' wages do not seem to benefit from the presence of CS and, further along the causal chain, we find no evidence that total household income improves with certification. The integrated synthesis of quantitative and qualitative studies shows that context matters substantially in all causal chains and multiple factors shape the effectiveness and causal mechanisms that link interventions associated with certification and the wellbeing of producers, workers and their families.

KEYWORDS: Certification; Private standards; Agriculture; Impact; Meta-analysis; Mixed-methods

^a Department of Development Studies, SOAS University of London, Thornhaugh Street, Russell Square, London WC1H 0XG, United Kingdom, co2@soas.ac.uk

^b School for International Development (DEV), University of East Anglia, Norwich Research Park, Norwich NR4 7TJ, United Kingdom

Highlights

- We find positive effects on prices and income from sale of certified products.
- However, we find no change in overall household income and assets.
- Wages for workers are not higher in certified production.
- Context is crucial in explaining variation in effects.
- Impact evaluations on certification need better study design and reporting.

1. Introduction

International agricultural trade has expanded rapidly in recent decades. As sources of agricultural products multiply in ever more complex supply chain systems, a growing set of new standards and regulations has come to affect the way agricultural commodities are produced, traded and consumed (Byerlee and Rueda 2015). A mixture of market, industry, relational and civic conventions increasingly shape the governance of value chains and the distribution of value therein (Auld et al. 2015). Exports from low- and middle-income countries (LMICs) in particular are increasingly covered by private voluntary standards that claim to certify the social and environmental sustainability of production conditions. This expansion of agricultural trade and associated standards may benefit producers and workers incorporated into global value chains, insofar as new conventions help empower these groups and improve control over the social and environmental effects of agricultural production. A rapidly growing empirical literature is seeking to evaluate whether and how the certification of private voluntary standards impacts the socio-economic wellbeing of producers and wage workers. This article presents the main findings of the first ever systematic review and meta-analysis of the literature on the effects of agricultural certification systems on direct producers and workers in LMICs. Our approach integrates quantitative and qualitative evidence on impact and causal mechanisms. We present evidence on key socio-economic outcomes and discuss the role of contextual factors in explaining these results.

Certification systems (CS)¹ are often multi-stakeholder initiatives with multiple drivers and shifting priorities. Historically, they frequently originated from, and were driven by, NGOs, as in the case of Fairtrade, a pioneering standards system that remains one of the most famous today (Raynolds 2017). However, other standards such as Utz Certified or RSPO (Roundtable for Sustainable Palm Oil) have been more business-driven. The main function of CS in agriculture, especially in the form of third-party certification, is to set voluntary standards with specific requirements for producers or suppliers, monitor their compliance (through independent auditors) and support producers to meet them, with the goal of

¹ The term standards system is also frequently used.

making agricultural production more economically, socially and environmentally sustainable and agricultural trade fairer to direct producers, i.e. farmers, and workers. Such schemes commonly, but not always, involve better performance and profitability through risk management, access to new and niche markets, and more predictable supply. Consumer-facing labels are usually used to communicate to the public that a product has been produced and sourced under specific standards and hence address both consumer and corporate public relations concerns (Aidenvironment 2017).

There is a voluminous literature about certification systems, their rationale, governance mechanisms, organisation of value chains, institutional features and how they shape the dynamics of markets of agricultural products and consumer behaviour (Raynolds & Greenfield 2015; Gibbon & Ponte 2005; Muradian, & Pelupessy 2005; Nelson and Pound 2009; Byerlee and Rueda 2015; Ruben 2012). However, this systematic review is more narrowly concerned with the evidence on the *impact* of these systems and their associated interventions on the *socio-economic wellbeing* of those who are supposed to be the ultimate beneficiaries of certification: *direct producers and hired workers*. A broad sweep of the abundant literature on certification, voluntary standards and their impacts on value chain participants in low- and middle-income countries (LMICs) suggests that the evidence regarding outcomes for producers and workers is inconclusive. Many studies report mixed findings or cases where effects are only marginal (Nelson and Martin 2013, Ruben 2012). Some conclude that CS may actually undermine the incomes of the poorest farmers (Henson and Jaffee 2008), some found effects only for richer farmers (Hansen and Trifković 2014), while others suggested CS help raise rural incomes and reduce poverty (Maertens and Swinnen 2009, Schuster and Maertens 2016). Various studies found negligible or even negative effects on employment conditions (Cramer et al. 2014a, Barrientos et al. 2003, Colen et al. 2012). Other studies reported positive impacts for some certification types, but not others (Chiputwa et al. 2015), or suggested that positive effects may dissipate due to over-certification (de Janvry et al. 2014).

Previous attempts to review and synthesise the evidence (International Trade Centre 2011, Blackman and Rivera 2010; Vagneron and Roquigny 2011) have shown that much of the existing body of empir-

ical literature is still characterised by evaluation designs vulnerable to validity threats, while the description of data collection and analysis tends to be poor, preventing assessments of the quality of the evidence (Terstappen et al. 2013, Cramer et al. 2014a, Ruben 2013). However, these reviews have important limitations with regard to the transparency of review process, the critical appraisal methods used and the approach to synthesis, which mean they cannot be considered systematic reviews as defined by Campbell Collaboration.² Many such reviews also focus only on selected CS – or even on a single system (e.g. Fairtrade in Terstappen et al. 2013, Nelson and Pound 2009, Darko et al. 2017). Therefore, a full systematic review, based around a statistical meta-analysis, was necessary to establish the state of the evidence on effects of certification on producers and workers, and the mediating factors that explain such effects or lack thereof. Such a review can also extract useful methodological lessons that may help improve the overall quality of the impact evidence generated by independent studies and research commissioned by CS.

Accordingly, this review set out to answer the following research questions:

1. *What are the effects of certification systems for sustainable agricultural production, and their associated interventions, on socio-economic outcomes for farmers, wage-labourers and their households?*
2. *Under what circumstances and why do certification systems for agricultural commodities have the intended and/or unintended effects? What are the barriers to and enablers of certification's intended and/or unintended effects?*

While the first question addresses the effectiveness question (i.e. do CS work?), the second question is critical in unravelling the causal mechanisms of impact and identifying barriers and enablers of CS effectiveness, taking into account how the context mediates between the implementation processes and the final outcomes (Weiss 1997, Pawson and Tilley 2004, White 2009). To answer the first of these research questions, hereafter RQ1, we conducted a statistical meta-analysis of the effect size estimates

² For more information on the Campbell Collaboration guidelines, see <https://www.campbellcollaboration.org/expectations-and-guidance-for-systematic-review.html>

provided by quantitative impact studies. For the second question, hereafter RQ2, we adopted a ‘thematic synthesis’ approach of the relevant qualitative evidence, as developed by Thomas and Harden (2008). The result is a theory-based, mixed-method systematic review that integrates both quantitative and qualitative evidence in order to provide combined answers not only on whether CS work, but also on how, why, when and for whom CS may or may not work.

The remainder of the paper is organized as follows. Section 2 lays out our analytical framework, provides a discussion of hypothesised causal chains linking certification to impact, and addresses the complications implied by the multiplicity of certification system interventions. Section 3 describes the methods used to identify and screen relevant studies, as well as our approach to data extraction, critical appraisal and data synthesis. Section 4 presents the main findings and integrates the results from quantitative (RQ1) and qualitative research (RQ2), with particular emphasis on the importance of context and how different kinds of contextual factors shape the effectiveness of certification. Section 5 concludes and presents suggestions for research and practice.

2. Framing the impact of certification systems

2.1. Standards, certification and interventions

Individual certification systems do not constitute a single distinct intervention. Rather, ‘certification’ is best seen as a bundle of interventions and associated organizational changes that varies across systems and contexts. These various bundles of interventions aim to have an impact, directly or indirectly, on different intermediate outcomes (e.g. prices, yields, farm revenues and wages) and endpoint outcomes (e.g. household income, health status and school attendance). Certification systems try to achieve their goals through a combination of standard-setting actions, compliance, capacity building and training for farmers and producers’ organisations, as well as different types of market interventions such as guaranteed market outlets, price premiums and credit facilities. Of course, some certification systems do not simply focus on the wellbeing of individual producers and workers upstream, but also aim to improve

their collective organization, and broadly the empowerment of vulnerable groups (Fairtrade 2015; Ronchi 2002; Lernoud et al. 2017). Furthermore, a focus on socio-economic effects on individual producers and workers may conceal the potentially significant organizational and institutional effects of establishing local certification bodies, local inspectors, certification officers, auditing processes, extension agents, and bringing new technologies to supply chains with the use of cellphone-based apps, computer database management, document production, etc. which were often absent in buyers and producers' organizations before certification bodies intervene. This is beyond the scope of the effectiveness review, but some of these issues are partly discussed in section 4.5 on contextual factors and the institutional 'environment' surrounding certification.

Regarding workers specifically, a wide range of CS in the past two decades have focused on the application of adequate labour standards relating to wages and non-wage working conditions, particularly with regards to health & safety. This is a result of two main trends. First is the growing awareness among consumers about the importance of labour standards, decent work and the prevention of the worst forms of exploitation in the production of widely traded commodities, partly driven by campaigning by international organizations like the ILO, NGOs like OXFAM or by media attention to notorious scandals like the Rana Plaza incident in the case of the global garment industry.³ Second, seeking to expand their reach, certification systems initially focused on small producers, have been extending their certification to large-scale plantations (Raynolds 2017). As a result most CS, even those focused on other dimensions of sustainability such as environmental protection, also include basic labour standards in their requirements. By and large these adhere to basic ILO guidelines such as respect for national legislation and minimum wages, as well as abolition of worst forms of work (forced labour or child

³ This is the incident at the Rana Plaza Centre, an eight-story commercial building, which collapsed in the Greater Dhaka Area, in Bangladesh, claiming the lives of 1,129 workers and many more injured. It became a turning point in the rise of labour standards in supply chains for the famous brands (Mezzadri 2015).

labour). However, many CS are also gradually engaged in deepening these standards by considering, for example, the application of a ‘living wage’.⁴

Certification systems have designed and implemented a wide variety of interventions across an even wider variety of settings. Each system includes different standards with multiple requirements. In some cases there may even be different grades of certification, as in the case of MPS for flowers, GlobalGAP and all of its different standards, or the different standards applied by Fairtrade to small producer organisations (SPOs) or hired labour organisations, i.e. large-scale plantations (HLOs). This poses an important challenge for evaluators, as the interaction of multiple objectives (e.g. medium and long-term socio-economic outcomes, environmental outcomes, etc.) with different standards and forms of interventions leads to a range of different pathways to socio-economic impact. This systematic review focuses on socio-economic effects for individual producers and workers, leaving out environmental outcomes and measures of empowerment at collective level. The aim was not to evaluate the work of all the different CS on all the outcomes they monitor (whether at collective or individual level), but rather to assess and synthesise the existing evidence on selected socio-economic outcomes associated with interventions under CS as defined in this review.⁵

2.2.Causal chains: towards a theory of change

Certification systems operate in complex environments at the sites of production and also have to be able to effectively communicate their actions and achievements to consumers, both of which require clear thinking about how impact is achieved. To organise their interventions, organisations working in the field of certification have recently begun articulating theories of change (ToCs). ToCs are structured process maps of the sequence of steps from interventions to outcomes that incorporate the assumptions needed for success at each stage of the causal chain (Vogel 2012). A growing number of CS (incl. Utz

⁴ ISEAL Alliance, the leading global network of social sustainability standards systems coordinates a current initiative involving seven certification systems in the estimation and enforcement of ‘living wages’ in their labour standards. See <https://www.isealalliance.org/get-involved/our-work/global-living-wage-coalition>

⁵ See Oya et al (2017) for an explanation of why social and environmental can be treated separately despite obvious linkages between the two sets of aims.

Certified, Rainforest Alliance and Fairtrade) have produced their own theories of change in order to guide the design of their systems, make impact evaluations theory-based and monitoring more structured around key pathways. Other organisations have at least produced sets of hypothesised causal links between inputs and impacts.⁶ For any theory-driven review a well-articulated theory of change is a useful tool for structuring the analysis. However, given the wide variety of intended outcomes and methods of intervention across different CS, there is no single theory of change that is valid for all types of CS. Still, attempts at ToC synthesis have been made. Nelson and Martin (2011, 2013) developed a ToC that is valid for multiple types of CS. And ISEAL, a membership organisation that represents several CS, has developed an impacts code to assist standards systems in their preparation of organisation-specific ToCs.⁷

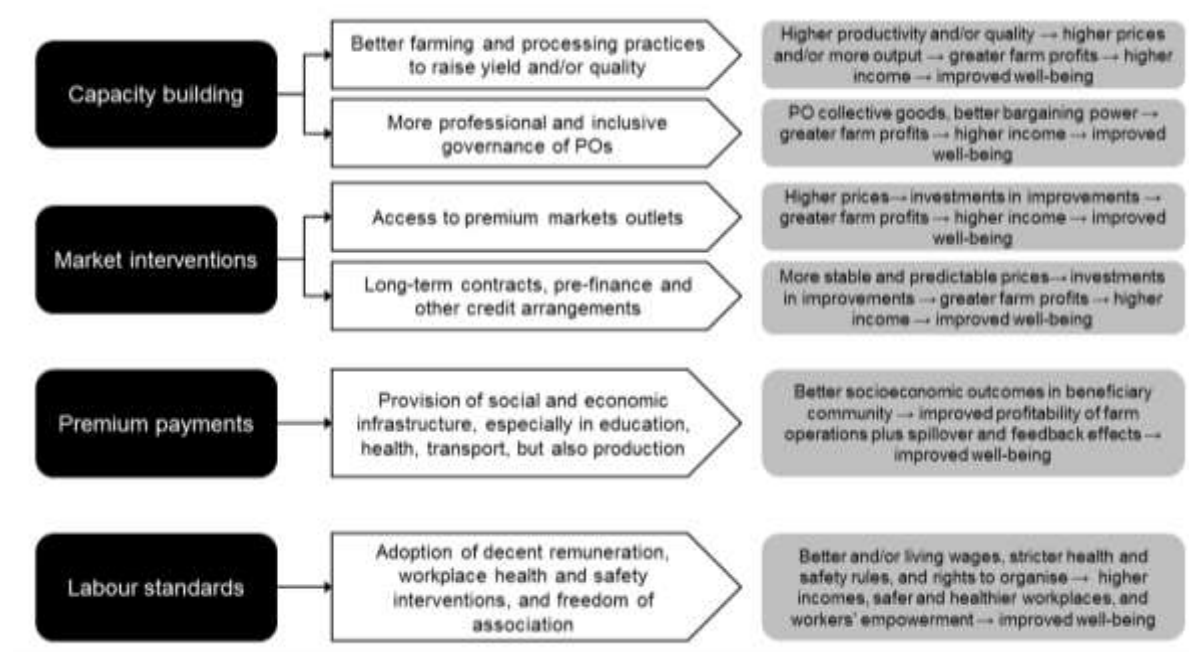
In order to organise the multiple systems, objectives and bundles of interventions under review into a coherent framework for the analysis of causal chains, we developed a simplified synthetic mapping of the causal chains involved by comparing the detailed ToCs produced by leading CS, the dimensions of certification proposed by the International Standards Map (ITC 2016) and the pathways to impact proposed by ISEAL. We group the most widespread interventions that come with the certification process around four blocks, each with common goals: (a) Capacity building through training and other forms of support to producers and their organisations to improve the sustainability, competitiveness and inclusivity of their production systems; (b) Market interventions (especially price measures and better contractual terms) to provide better incentives and market conditions to improve the welfare of producers and their production standards; (c) Additional payments for social and business investments (frequently referred to as a premium), usually provided to POs, to enhance socio-economic outcomes and the strength of POs; and (d) Labour standards for decent working conditions across different kinds of employers. Each of these categories includes a range of specific interventions. These impact pathways

⁶ See, for example, Fairtrade (2015), and Utz (2017), which has now merged with Rainforest Alliance, as examples of ToCs developed in compliance with ISEAL Impacts Code v2. ISEAL (2016: 6) also provides a list of CS with links to their ToC or similar attempts at linking inputs and impact.

⁷ For a list of ISEAL member organisations see, <https://www.isealalliance.org/our-members>.

are relevant to most standards systems in agriculture. Figure 1 summarises these hypothetical causal chains.

Figure 1 - Causal chains linking intervention to impact



This framework must contend with a number of complications. Although it is conceptually helpful to separate the different sets of interventions into their own causal chains, in practice it is likely that causal chains intersect and interact, especially when multiple interventions take place at the same time, as in the case of price incentives and capacity building supplied by Fairtrade, for example. A further complication arises due to the difference between intermediate and endpoint (or final) outcomes, and the length of some of the causal chains involved. Intermediate outcomes are results of the intervention, but themselves contribute to endpoint outcomes that are the final aims of the certification interventions. Examples of the former include higher prices or increased yields, while higher household income and reduced poverty are examples of the latter. There is limited evidence to support many of the impact pathways directly linking certification interventions to final outcomes such as household income and poverty status, an important reason being that ‘certified crops only represent a – sometimes small – share of land and revenue streams’ of certified producers (Ruben 2017: 87). This is a major challenge for impact evaluations, which often do not collect enough evidence on the proportion of sales that go to certified

channels, both at individual and collective levels. Moreover, as each certification combines a number of different interventions, it is difficult to tell whether an estimated effect suggests that a particular type of intervention is effective or not. Most impact evaluations can only say whether being part of a certification system has any impact on the outcome. It is therefore difficult to ascertain to what extent any particular endpoint outcome is associated with specific interventions. It is easier to show what impact interventions have on a range of intermediate outcomes, such as participation in training, access to support services, adoption of improved practices, changes in crop choices and producer prices. Indeed, most ToCs developed by CS emphasise the increasing importance of context as we move along the causal chain from interventions, through intermediate outcomes, to final outcomes. By way of illustration, we discuss one type of intervention – good agricultural practices – in greater detail and lay out its hypothetical linkages with endpoint outcomes.

Interventions designed to introduce good agricultural practices, or GAP, are very common across a wide range of CS (e.g. Utz-Rainforest Alliance or GlobalG.A.P.).⁸ They typically include: training and extension services for better farming practices to improve quality, productivity and/or food safety; assistance for professional farm management; training for PO management, especially in relation to coordination between members, exporters and buyers; and product quality checks. The pathway assumes that farmers adopt improved practices after training, and that such practices lead to a higher value of output. Additional assumptions are made about the availability of the required services and inputs, especially credit, to certified farmers. If these assumptions hold and interventions are implemented as planned, then intermediate outcomes are expected to materialize. Improved farm management, through behavioural or attitudinal changes, combined with more resources should lead to investments which improve the quality and/or productivity of traded commodities. Higher productivity and better quality should in turn increase the value of farm output. Access to more remunerative ‘niche’ markets for certified goods, coupled with more efficient production, increases profitability on a sustained basis. Increased profits

⁸ At the time this review was completed, Utz and Rainforest Alliance were separate standards and organisations. They merged in 2017 given their affinities in aims and practices. See <https://utz.org/merger/>

lead to higher farming incomes and overall household income, so long as the growth in certified production is not offset by abandoning other profitable income generating activities. Sustained increases in household income may in turn affect wealth and household investments in education and health.

Our quantitative synthesis looked at outcomes relevant to all four of the causal chains presented in Figure 1, and we use qualitative data to explore the validity of assumptions underlying these hypothetical chains. We present the results of the meta-analysis for all relevant outcomes for which we have a reasonable number of studies in order to draw meaningful conclusions. We then combine our quantitative and qualitative syntheses along three key causal chains, namely income from certified production, wages, and household income, which include all the relevant reported outcomes.⁹ The outcomes presented in the most detail are those which we value for conceptual reasons, and which are more central to claims generally made by certification systems about their impact on socio-economic sustainability. These tend to centre on improvements in income, decent working conditions, and improved living conditions as key components of ‘sustainable livelihoods’, as described by Fairtrade in their ToC (Fairtrade 2015: 7). Each of these three key outcomes is linked to more than one of the causal chains described in Figure 1. Income from certified production is linked to market interventions and capacity building. Market interventions directly seek to secure higher prices for certified goods with the hope that higher prices translate into higher incomes, while capacity building interventions aim to enhance product quantity and quality, which also raise incomes. Wages are connected to the setting and enforcement of labour standards, and of course to interventions which directly or indirectly attempt to raise wage levels, including empowerment through trade unions or other workers’ organizations. But workers may also benefit from premium payments that help develop local socio-economic infrastructure and services. Lastly, of the outcomes discussed here household income is furthest along the causal chain, and can therefore be affected by each of the other outcomes. It is partly dependent on income from certified production, and so is tied to market interventions and capacity building, but it is also influenced by local spillover effects from investments funded by premium payments. As household income is the sum of

⁹ Synthesized evidence on producer prices and yields is incorporated in the discussion of the causal chain on certified farm income.

all income generating activities by all household members and some household members may be wage workers, it can also be influenced by interventions targeting labour standards. Unintended effects may include the diversion from other income generating activities because of higher labour demands associated with certification.

3. Review methods

Before we present the findings of the review we turn now to a discussion of the review process itself. We conducted this review in line with the Campbell Collaboration guidelines for systematic reviews in the social sciences, which demand that all steps in the review are documented and made transparent. Accordingly, we have published a pre-review protocol laying out search methods, study inclusion criteria and procedures for data synthesis (Oya et al. 2015), as well as technical report detailing all review methods (Oya et al. 2017).

3.1. Scope

To be considered for inclusion studies had to present primary evidence on the socio-economic impact of certification systems for agricultural production on direct producers, i.e. farmers and workers, in low- and middle-income countries, or present evidence on factors enabling or inhibiting such impact. We only considered certification systems that have clearly defined socio-economic goals and third party auditing. Studies had to be published no earlier than 1990 in English, French, German, Spanish or Portuguese. Eligible impact studies had to report on at least one of the following socio-economic outcomes: household income or consumption, health, education, gender equity in the above outcomes, individual empowerment, returns to certified production, yields, price levels for certified products, wages, non-wage working conditions, collective empowerment of producers' and workers' organisations, or investments in services and infrastructure.

For RQ1 we only included controlled quasi-experimental studies, as well as *ex post* observational studies with non-treated comparison groups. Impact studies had to feature research designs and analytical methods that take into account selection bias and other confounding factors. Such designs include

matched comparisons, difference-in-difference analysis, instrumental variables, and a variety of multi-variate regression techniques. We would have included randomized controlled trials (RCTs) but no such studies were available at the time of the review.¹⁰

For RQ2 we included studies that contained relevant and substantive qualitative evidence, regardless of the study design. This means that evidence could be provided by studies exclusively using qualitative methods, but also by quantitative or mixed-methods impact evaluations, as long as the evidence could serve to illuminate issues of implementation, distribution, barriers and facilitators to impact.

3.2. Search

We searched 14 relevant social scientific databases using specifically designed search strings.¹¹ An example search string can be found in Appendix 1. As the literature on certification systems includes large amounts of material that is not indexed in academic databases, we also conducted extensive ‘manual’ searches of more specialist databases and of the websites of certification bodies, NGOs and research organisations. Finally, we hand-searched non-English databases and websites. These hand-searches covered 40 databases and organisations.¹² The last string search was conducted in October 2015 and the last hand searches were undertaken in December 2015. A few items were added after consultation with our study’s advisory committee in July 2016.

3.3. Screening

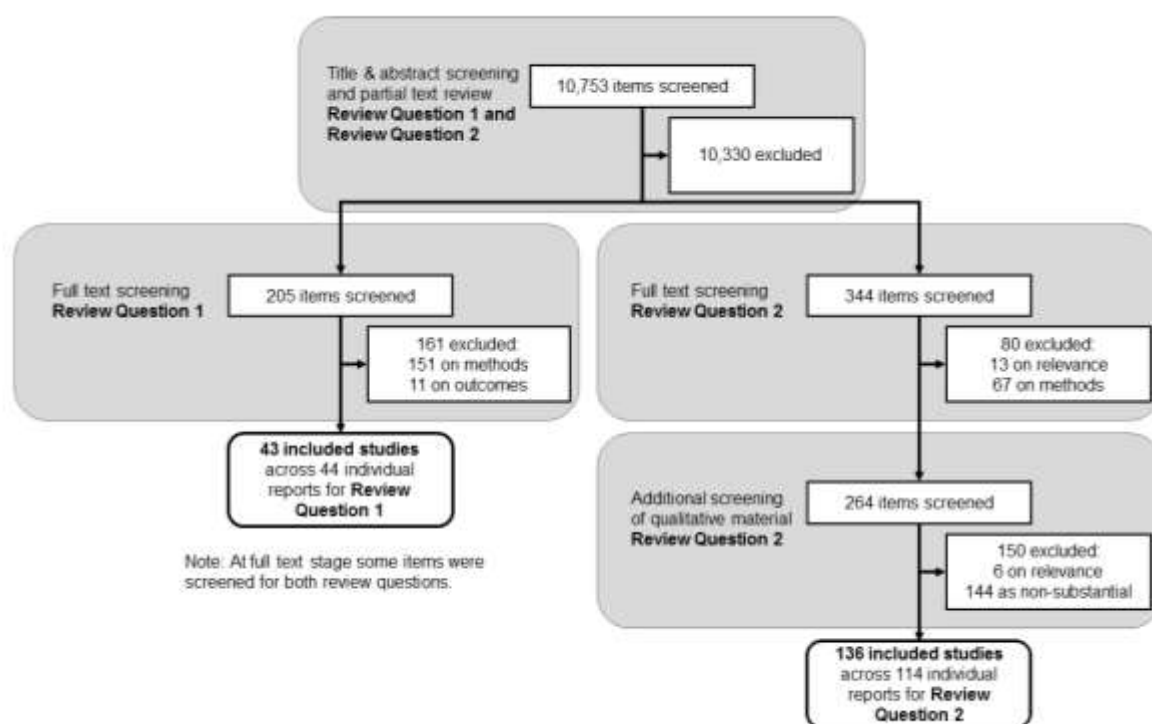
¹⁰ While two RCTs on certification were in progress at the time of writing, neither was completed in time for inclusion in the review.)

¹¹ These were: AgEcon, Africa Wide, CAB Abstracts, International Bibliography of the Social Sciences (IBSS), Social Sciences Citation Index (SSCI) / Web of Science, EconLit, US National Agricultural Library/AGRICOLA, JOLIS (WB/IMF), British Library for Development Studies, IDEAS RePEc, International Institute for Environment and Development, 3ie Systematic Reviews and Impact Evaluations Database, ILO Labordoc, and the Campbell Library.

¹² Please see Oya et al. (2017) for the full list.

All studies identified during the search process were screened against clearly defined inclusion criteria and any study that did not meet these criteria was excluded from the review. As our two research questions require very different types of data, we developed separate inclusion criteria for each. We used a multi-stage screening processes, moving from screening on titles and abstracts, through a partial text review, to a review of the full study text, with increasingly stringent criteria at each stage. At the full text stage studies were independently double-coded and any disagreements were moderated by the principal investigator. At this stage we also applied a final check against methodological inclusion criteria and made a last assessment as to study relevance. This yielded a total of 43 studies across 44 different reports for RQ1, and 136 included studies across 114 reports for RQ2.¹³ A detailed list of all studies included in RQ1 can be found in Appendix 2, while the screening process is summarised in *Figure 2*.

Figure 2 - Summary of the screening process without intermediary stages



¹³ As is common, we make a distinction between a study, by which we mean a unique dataset, and a report, which is a single publication. The results of studies are frequently written up in more than one report. Conversely, a single report may contain findings from a number of studies.

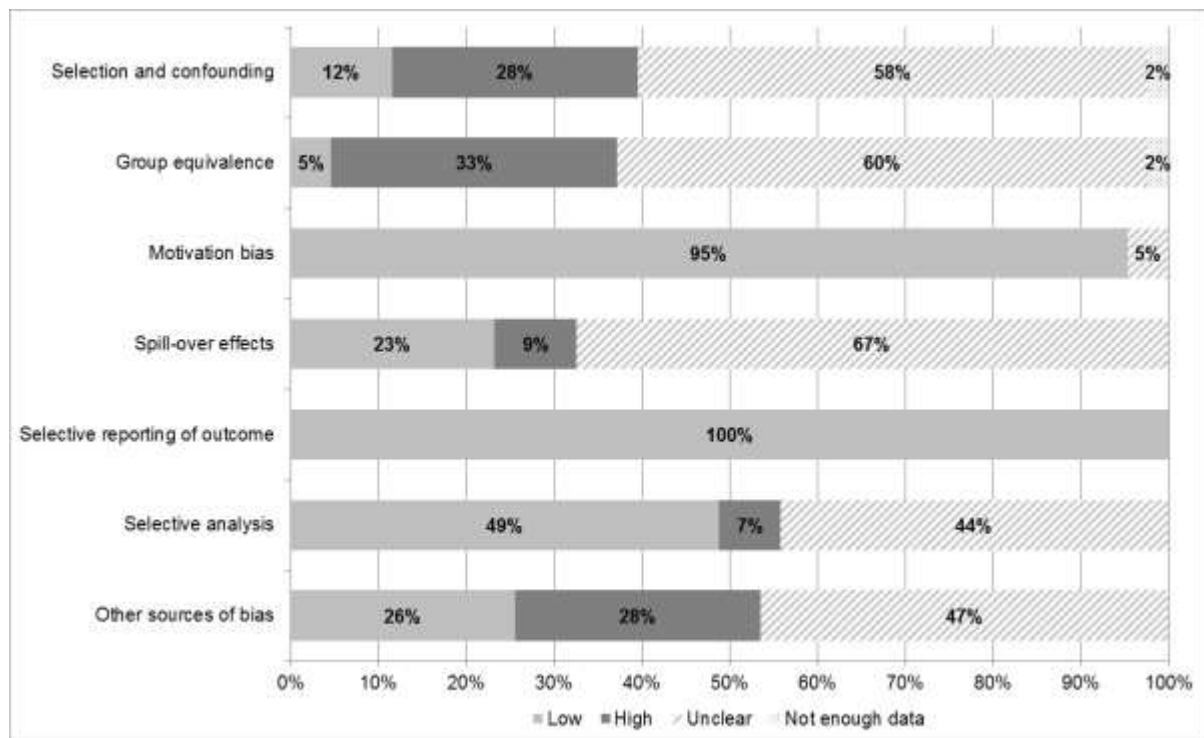
3.4.Data extraction and critical appraisal

Data was extracted from all included studies. Across all studies we extracted bibliographic information, data on project type and implementation, and methodological information. For RQ1 studies we also extracted the data needed to calculate effect size estimates and standard errors, as well as for moderator and sensitivity analysis. For RQ2 studies we extracted information related to implementation, distributional dynamics, and barriers and facilitators.

Prior to the data synthesis each included study was critically appraised in terms of its methodological quality. For quantitative studies the most important aspect was the degree to which the study design took account of selection bias, how the analysis controlled for other confounding factors, and whether equivalence between treatment and control groups was ensured. Figure 3 shows the risk of bias rating for RQ1 studies across the different categories of assessments. Many studies fare poorly in the most important areas, i.e. taking account of selection bias and confounding factors and ensuring group equivalence. The high incidence of ‘unclear’ ratings demonstrates that many studies fail to report important information on study design, implementation and data analysis, a weakness that needs to be corrected in on-going and future impact evaluations on certification. For RQ1 studies we produced overall risk of bias ratings for each included study. Just two studies (5% of included studies) received a low risk of bias rating, while more than half were rated as high or critical risk of bias.¹⁴

¹⁴ Studies that received a critical risk of bias rating were retained in the review, but were excluded from the statistical meta-analysis.

Figure 3 - Risk of bias rating of quantitative studies (RQ1) by category of bias

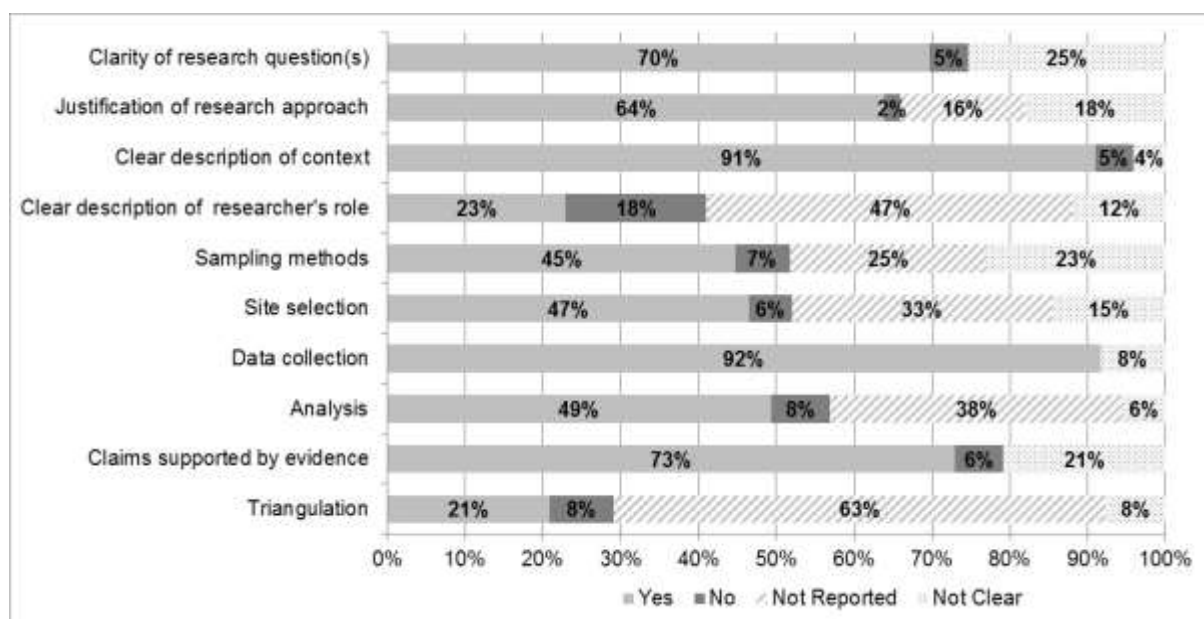


For RQ2 conducting the critical appraisal was less straight-forward, as no consensus exists regarding how to assess the quality of qualitative evidence (Campbell et al. 2003, Atkins et al. 2008, Thomas and Harden 2008), or even whether it should be assessed at all (Schwandt 1996). Although some authors recognise the need to define what constitutes ‘good evidence’ before proceeding to the synthesis (Popay et al. 1998, Dixon-Woods et al. 2004, Hannes 2011), concerns about excluding studies on the basis of quality assessments remain, mainly due to fears of excluding good research due to inadequate appraisal criteria that not only lack consensus on what counts as good evidence but also the ability to do justice to the variety of qualitative approaches (Sandelowski et al. 1997, Dixon-Woods et al. 2004, Atkins et al. 2008).

For these reasons we did not produce an overall quality rating for RQ2 studies, nor did we exclude any study that already met the inclusion criteria on the basis of the appraisal. However, we assessed whether studies had defined research questions, justified their chosen approach and selection of research site, and whether they provided clear descriptions of the context and the researcher’s role. We also considered whether data collection methods were adequate, whether the methods of analysis were appropriate,

and whether the claims made were supported by the evidence presented and based on triangulated information. Figure 4 shows the critical appraisal results for RQ2 studies. The biggest weaknesses were a lack of triangulation, as well as a lack of engagement with the positionality of the researcher. The latter is especially problematic in a field where a lot of researchers have ties to certification systems as specialists in specific CS or as researchers doing commissioned work for these systems. Main strengths were the description of context and account of data collection methods.

Figure 4 - Critical appraisal results for qualitative studies (RQ2) by category of assessment



3.5.Synthesis

The main aim of any systematic review is to synthesise the information provided by the included studies. As this is a mixed-methods systematic review, we first conducted a meta-analytical synthesis for RQ1 and a thematic synthesis of qualitative evidence for RQ2 before producing an integrated synthesis in which the statistical results for RQ1 were put in dialogue with the results of the qualitative synthesis.

A statistical meta-analysis is a method for producing a weighted average effect size estimate for a given outcome across a number of studies. We calculated effect size estimates, for which we used standardised mean differences (SMDs), and associated confidence intervals for all outcomes in all included studies. We then synthesised the data using inverse variance-weighted random effects meta-analysis (Borenstein

et al. 2009; Lipsey & Wilson 2001). Study weights are largely a function of standard errors, where larger standard errors decrease study weight. So studies that estimate effects more precisely contribute more to the pooled effect.

Prior to synthesis we applied small-sample bias correction using Hedges' method. SMDs and their standard errors were calculated using Wilson's effect size calculator, which uses standard formulae provided in Lipsey & Wilson (2001).¹⁵ Not all studies provided the information necessary to calculate effect sizes. For instance, the pooled standard deviation is not commonly reported. In these cases we used the standard deviation for the control group, or relied on approximations for the SMD, as provided by Waddington (2014).

We calculated our results both with and without correction for potential unit of analysis errors, but found little difference between both sets of results. We examined how results vary by subgroups of studies, but, given the limited amount of data available, we did not conduct any meta-regressions. To assess the likelihood of publication bias we conducted sensitivity analysis comparing peer-reviewed and non-peer-reviewed studies, and used funnel plots and Egger's test for funnel plot asymmetry to assess the possibility of small study bias in the included studies. The intuition behind such bias is that smaller studies, which generally are less precisely estimated, tend to produce more significant effects. This can lead to a preponderance of positive findings over null and negative findings in published small studies, as the latter are harder to publish. We find that non-peer-reviewed studies generally report larger effect size estimates than peer-reviewed studies, but the direction of the effect remains unchanged. Egger's test indicates the possibility of publication bias only in the case of the result on wages ($p=0.022$ in the test against the null hypothesis of no small study bias). However, we do not believe that this result indicates publication bias. Observed heterogeneity and plot asymmetry may be due to real substantive differences across studies and interventions (Sterne et al, 2011). More importantly, the asymmetry in this instance is due to negative results, whereas publication bias normally implies a relative absence of

¹⁵ The calculator can be found at <http://www.campbellcollaboration.org/this-is-a-web-based-effect-size-calculator/>.

negative and null findings. If publication bias were to blame we would have to accept the rather unlikely thesis that statistically significant positive results are not being published.

To synthesise the data for RQ2 we used an iterative process to produce a thematic synthesis following Thomas and Harden (2008). Under each of the three core themes of the analysis – implementation, distribution, barriers and facilitators – sub-themes were developed *a priori* while emerging themes were added to during the coding process. Once this coding structure had been developed, detailed descriptive themes were developed from the coded material. These were used to generate analytical themes, which extend the analysis beyond the content of any single study. To minimise the impact of individual biases all analytical themes were discussed by three team members. Finally, we produced an integrated synthesis that seeks to explain and give context to our statistical results by drawing on the analytical themes we developed.

4. Results and discussion

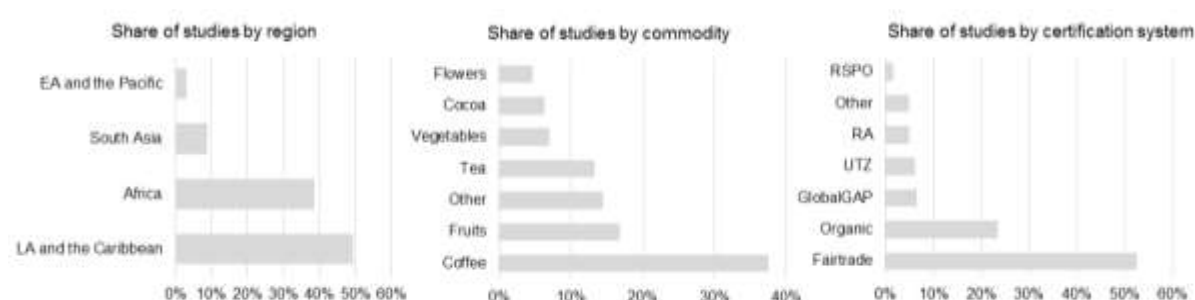
4.1. Settings and pooled effects

This section presents our findings, integrating the key results from both the meta-analysis of quantitative impact evaluations and the qualitative research synthesis. In other words, we use the main lessons from qualitative studies to help us understand quantitative effectiveness with sensitivity to context and heterogeneity. An important challenge for such an integrated synthesis is the scarcity of ‘linked’ studies, i.e. of studies that contribute to both RQ1 and RQ2. Impact evaluation studies that provide sufficient evidence to answer qualitative questions alongside the effectiveness question are scarce in the certification literature, while relevant qualitative material from separate studies conducted in the exact same settings as contemporaneous impact evaluations is also hard to find. In our case, only seventeen studies contributed to both the quantitative and qualitative synthesis, a fact that reflects a wider shortage of high-quality mixed methods impact evaluation studies.

We concentrated on outcomes that are shared across certification systems striving to achieve socio-economic sustainability. Our focus is on endpoint outcomes, i.e. those further down the causal chain,

such as household income, wealth (assets), health, and education, as well as on intermediate outcomes, which are more directly linked the interventions undertaken by certification systems. Certification interventions may or may not have a direct impact on endpoint outcomes, and the effects depend on a range of contextual factors, as well as on unintended consequences. Intermediate outcomes include (farm-gate) prices, product yields, wages and – a little further along the chain – income generated from the sale of certified produce.¹⁶ As discussed, the causal pathways to impact for certification systems rely on a number of assumptions. Hence, the synthesis of qualitative research, and particularly relevant ethnographic work, is critical to grasping the role of contextual and other factors that act as barriers and facilitators to impact.

Figure 5 - Distribution of included studies by geography, commodity type, and certification system

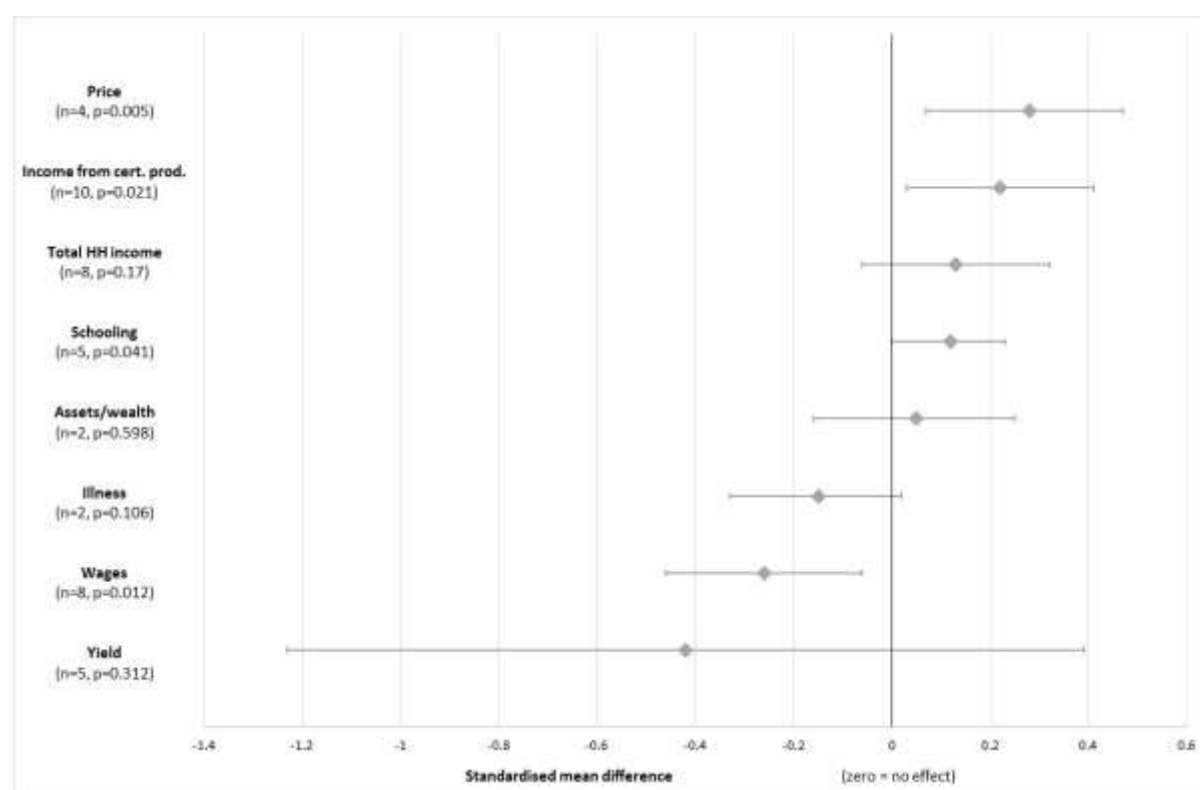


Before presenting the results for selected outcomes, it is necessary to provide some background information on the settings, value chains and types of standards systems that are covered by the available evidence. As Figure 5 illustrates, the studies included in this review are heavily skewed towards certain contexts, commodities and systems. Most of the included studies cover settings in Latin America and the Caribbean (49%) or Africa (39%), while comparatively few studies discuss South Asia or East Asia and the Pacific. Similarly, 38% of studies deal with coffee production, while the next most researched commodities, fruit (17%) and tea (13%), receive much less attention. The situation is the most pronounced with regard to certification systems. The available literature tends to focus on a limited set of

¹⁶ Measures of income from certified production compare the income producers receive from the production and sale of a particular certified commodity, such as coffee, with the income received by otherwise equivalent groups producing the same commodity but lacking certification. Income from certified production is not the same as farm income. While income from certified production is part of a household's farm income, that household may also produce non-certified goods, and so income from certified production may only be part of a farm income portfolio.

CS, and just one system, Fairtrade, accounts for over half of all included studies (52%), despite the fact that there are many other CS and some with more outreach than Fairtrade. This dominance of Fairtrade in the existing literature is also one of the reasons why certain regions and crops have been more frequently studied than others: Fairtrade is very active in Latin America and Africa, and in the coffee, tea and banana sectors.

Figure 5 - Overview of pooled effects and their 95% confidence intervals for all outcomes



We turn now to the results of the meta-analysis. Figure 5 summarises the results of the quantitative synthesis, based on 43 studies. For each outcome we report our estimate of the pooled effect size estimate (as SMD), along with a 95% confidence interval, the p-value of the significance test for the pooled effect and the number of studies used to estimate each pooled effect.¹⁷ Overall, the available evidence does not paint a clear picture of the impact of certification systems, as there is a range of positive and negative effects across outcomes and several cases of outcomes for which the synthesised effect is not

¹⁷ We considered a pooled effect to be statistically significant if $p < 0.05$.

statistically significant. The precision with which the pooled effect can be estimated varies widely from outcome to outcome. Starting from outcomes more directly linked to CS interventions, we found an average reduction in yields or output per hectare (SMD -0.42, 95% CI from -1.23 to 0.39) for certified farmers, although the effect is not statistically significant ($p=0.312$), the standard error is large, and only five studies could be included. Based on only four studies, the effects on prices is positive (SMD 0.28, 95% CI from 0.06 to 0.47), and the effect is statistically significant ($p=0.005$). We synthesized the effects of ten studies on income from certified production, which show a statistically significant increase in income from certified production (SMD 0.22, 95% CI from 0.03 to 0.41, $p=0.021$). The picture for wages is much less promising. On the basis of eight studies the synthesised effect on wages is negative (SMD -0.26, 95% CI from -0.46 to -0.06) and statistically significant ($p=0.012$).¹⁸ Possibly one of the most practically important endpoint outcomes for farmers is the change in total household income as a result of certification. In this case, eight studies show a combined increase in total household income as a result from certification (SMD 0.13, 95% CI from -0.06 to 0.32). The pooled effect is however not statistically significant ($p=0.17$), reflecting the substantial variation of effects across different studies. We have the most limited evidence (in terms of number of studies) for effects on the wealth (measured as assets owned) and health of producers. The pooled effect size estimates point to a reduction in illness and an increase in control over assets, though neither effects is statistically distinguishable from zero, which is not surprising as both findings rely on a very limited number of studies and these studies have diverging results. Finally, the synthesis of effects for schooling, based on only five studies, suggests a statistically significant improvement in school attendance among certified producers (SMD 0.12, 95% CI from 0.00 to 0.24, $p=0.041$).

We now move to a more detailed discussion of the integrated results of the quantitative and qualitative syntheses, considering effects on income from certified production (which also reflect the effects on prices and yields), wages for labourers, and total household income. For each of these we provide forest plots detailing the effect size estimates contributed by each study, along with their confidence intervals,

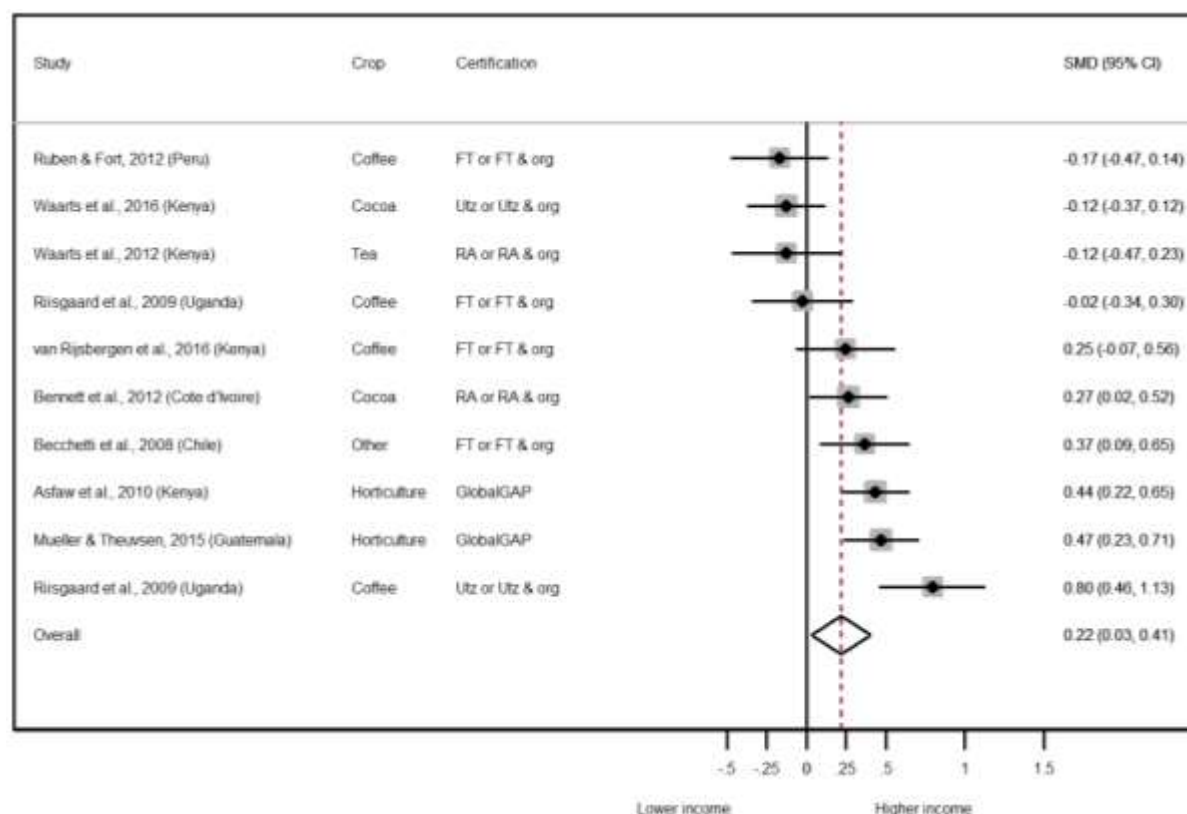
¹⁸ However, this result is largely driven by four studies conducted by the same research team (Cramer et al. 2014a).

the geographical setting of the study, the commodity and certification system under study, as well as the overall effect and confidence interval (represented by a diamond). Grey boxes around the effect size estimate indicate the weight assigned to each study in calculating the pooled effect, where larger boxes represent a greater weight.

4.2. Effects on income from certified production

Increased income derived from certified production should be expected to follow from interventions that aim to improve the profitability and trading relations for producer beneficiaries compared to those who are not part of a certified supply chain. As just discussed, the results of the quantitative synthesis for incomes from certified production (SMD 0.22, 95% CI from 0.03 to 0.41) are extracted from 10 studies, the largest number of all considered outcomes. The pooled effect masks some important differences between the individual studies that contributed to this result. Six studies produced positive and generally statistically significant effects, while four studies found negative or no effect, see Figure 6. Half of the studies were rated as having moderate risk of bias, and the other remaining ones were rated as high risk of bias, with only one exception as having low risk of bias (van Rijsbergen et al. 2016). When we pool the results by risk of bias group we find that each set of high (Waarts et al. 2012, Waarts et al. 2016, Becchetti et al. 2008, Asfaw et al. 2010), moderate (Ruben and Fort 2012, Riisgaard et al. 2009, Bennett et al. 2012, Mueller and Theuvsen 2015) and low risk of bias (van Rijsbergen et al. 2016) produces positive but statistically insignificant effects.

Figure 6 - Forest plot for income from certified production



Although the overall effect is positive, there is substantial variation in the underlying study-level effects. As the individual studies cover a range of different CS, as well as geographical areas and product types, these results suggest that the impact of CS may depend, in part, on circumstances. Income from certified production depends on the amount produced (i.e. yield), and the price per unit of output received by the producer. Subgroup analysis by type of CS returns positive but statistically non-significant pooled effects for studies covering Fairtrade (SMD 0.11, 95% CI from -0.14 to 0.36), Rainforest Alliance (SMD 0.09, 95% CI from -0.29 to 0.48) and Utz Certified (SMD 0.33, 95% CI from -0.57 to 1.23), while studies covering GlobalG.A.P. yield a pooled effect that is both positive and statistically significant (SMD 0.45, 95% CI from 0.29 to 0.61). The CS that do better according to the limited number of studies found (i.e. GlobalGAP) would suggest that a combination of more effective capacity building through ‘good agricultural practices’ (GAP) for productivity increases, combined with more remunerative markets, is more effective. By comparison, the evidence from the few studies on Fairtrade and Rainforest Alliance is less clear.

Indeed, evidence on yields suggests not all CS target productivity increases. For example, despite the overall aim to expand the aggregate volume of certified sales, increased individual farmer output is not a target outcome for Fairtrade, which focuses on market relations and the governance of producer groups. Moreover, Fairtrade is often applied alongside organic certification which may be associated with lower average physical yields in contexts of low-input agriculture, due to stricter quality criteria as well as the transition process from conventional to organic production that can lead to product losses (Nelson and Martin 2013, Shreck 2002, Abarca-Orozco 2015). According to our synthesis the overall effect of CS on yields is negative but not statistically significant. Subgroup analysis of the yield effects between CS corroborate the expectation that those explicitly targeting production efficiency, such as UTZ and Rainforest Alliance, obtain positive or at least non-negative results, whereas the few studies reporting on Fairtrade settings find negative yield effects. Therefore, the limited quantitative evidence we have suggests that yield improvements only make a marginal contribution to the overall effect on income from certified production.

The price linkage appears to be a more important contributor to the effect on income from certified production. According to our quantitative synthesis, based on only four studies, certification is associated with higher prices (SMD 0.28, 95% CI from 0.06 to 0.47). The positive effect is associated with different kinds of CS, including Fairtrade, GlobalG.A.P. and several other systems, but the size of the price effect is largest for GlobalG.A.P. These results indicate that markets may be more remunerative partly because of the type of commodity (higher value-added horticulture versus coffee or cocoa) and partly because of the quality demands associated with these CS. In other words, the type of product and associated supply chain also matters. However, the results on income from certified production suggest that higher prices are not always high enough to compensate for lower yields.

The synthesis of qualitative research has shed some light on the possible barriers and facilitators that underpin the quantitative effects described above. Due to space limitations we focus on the most salient aspects here. The effects on prices seem largely affected by the type of price mechanism. CS differ in the systems they use to support the prices paid to producers. The most common are paying a price premium (as applied by Fairtrade and Utz), instituting a price floor (Fairtrade) or simply relying on

quality-related prices for higher-end markets (e.g. GlobalGAP or MPS). Floor price interventions are mostly effective during commodity price slumps, for instance in coffee, a highly volatile commodity (Dowdall 2012, Milford 2014, Valkila 2009). One important offsetting barrier is the cost of certification which can negatively affect adoption and the effectiveness of price incentives, especially when they are combined with organic standards, which can significantly increase production costs, mainly through higher labour costs (Milford 2014, Jaffee, 2006, Abarca-Orozco, 2015). In addition to the *direct* certification costs (which are meant to help producers meet registration and auditing costs), new farming practices stemming from capacity building interventions often also entail additional investments in labour or inputs, whose cost may not be affordable to all producers. When such practices are designed to improve productivity, there may be a trade-off between long-term yield gains and short-term increases in costs and workload (e.g. Arce 2009, Lyon, 2005, Cofre et al. 2012). One limitation of the available evidence is that these trade-offs over different timeframes are not be adequately captured.

Even when certified prices are higher, the effect on income from certified production is limited by the extent to which markets are able to absorb the total volume of certified products. Many qualitative studies stress the extent to which certified POs and plantations only sell a fraction of their output, sometimes quite a low fraction, to certified channels (e.g. Beuchelt 2009, Dragusanu and Nunn 2014, Nelson and Martin 2013, Ronchi 2002, Valkila and Nygren 2008, Minten et al. 2015). In these cases the overall size of the certification rent is limited, but it has to be distributed among large numbers of individual producers. Positive price effects from certification may dissipate in producer organisations with large numbers of members that sell only a limited proportion of output as certified.

The structures of different value chains, and the relations between actors therein, also affect the impact of interventions on certified income. Qualitative research suggests that effects on certified income are magnified when communication between producers and buyers is effective, standards are properly applied and additional buyer support outside the certification channel is provided, especially in the form

of pre-finance, long-term contracts and assistance with product quality/grade improvements ¹⁹(Jena et al. 2012, Subervie and Vagneron 2013, Valkila and Nygren 2008). Thus an important question is whether and to which extent certified trading relations replace or improve pre-existing trading relations. Qualitative evidence indicates that the producer-buyer relationship may not be substantially altered (nor further improved) by the introduction of certification, especially when pre-existing market relations are deep-rooted and personal trust and loyalty are strong. Many studies show certification generally does improve access to export markets (e.g. Stathers and Gathuthi 2013, Aidenvironment 2016, Trauger 2014), so when these new markets come with improved trading relations for producers the effect on certified income is clearer. However, there are also cases reported where non-certified buyers deemed CS not sufficiently effective and offer direct benefits to their suppliers without certification. These were tailored to producers through personalised non-certified channels and often superior to those offered by certified markets, as is the case of “relationship” coffee (Valkila and Nygren 2008).

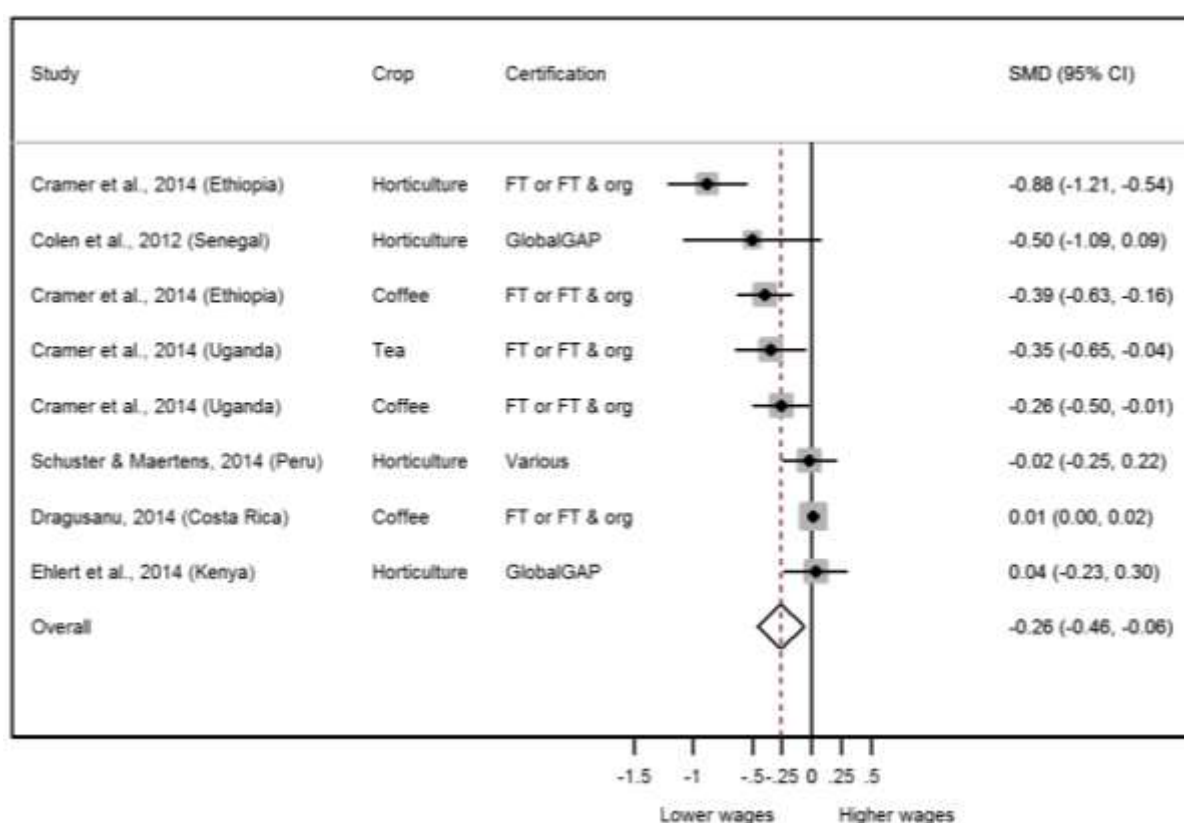
The effect on certified income is also dependent on the use of the different kinds of ‘premium’ and whether they affect farm profitability. In the case of POs the distribution of such rents is also important. Thus the translation of collective-level effects (e.g. benefits from PO-lead investments using the certification premium, as in Fairtrade) into outcomes for individual producers depends on what kinds of investments have been privileged, which will determine what kinds of producers benefit most. The choice of investments at the PO level in turn is influenced by the specific configurations of power relations affecting the management of POs. Several studies provide substantial evidence on premium use and distribution (e.g. CESU 2012, Cepeda et al. 2013, Riisgaard et al. 2009, van Rijn et al. 2016), and instances of elite capture in the management and use of ‘social premium’ (e.g. Sutton 2014, Staib 2012, Dolan 2010).

4.3. Effects on wages

¹⁹ Improvements in the quality of the product, especially those that carry differential grades for export markets (e.g. coffee), can carry a substantial price premium.

Turning from farmers to agricultural wage workers, we find, on the basis of eight studies, that wages for workers engaged in certified production were lower than for workers working for uncertified employers (SMD -0.26, 95% CI from -0.46 to -0.06, $p=0.012$). Of the eight studies, all but three provide negative effect estimates and there were no studies reporting statistically significant positive effects on wages, see Figure 7. However, all statistically significant effects come from the four studies in Cramer et al (2014a), and excluding these four studies returns a pooled effect that is not statistically distinguishable from zero. The literature on wage employment is much more limited than the literature on effects on producers. Of all the studies included in this review just 23% analyse the effects of CS on workers, while the rest deal with farmers. Most CS share the adoption of labour standards, following either International Labour Organization conventions or more demanding requirements, as their most important intervention with regard to workers. The different effects may lie in the extent to which the specific requirements exceed the standards established by the ILO or indeed the minimum legal standards found in specific national contexts.

Figure 7 - Forest plot for wages



One problem is that CS do not target all workers. There is a large-scale bias in the application of standards and the evaluation of their effects, which mostly focus on agribusiness. This means that large segments of the employed agricultural population may be ‘invisible’ to CS interventions and to evaluators, primarily because labour standards do not seem to be consistently applied to workers employed by small farmers (Cramer et al. 2014a, Nelson et al. 2013, Waarts et al. 2016,²⁰). One reason could be that both CS and researchers assume that conventional labour standards may be too demanding in these cases, as smallholders are too poor and subject to volatile market conditions to offer decent work standards to their seasonal and casual workers (Shreck 2002, Valkila 2009, Heller 2010, Valkila and Nygren 2008). Cramer et al (2014a), whose studies also include quantitative effects on wages for different kinds of employers, suggest scale is not the only relevant independent variable, since lower wages are also found in certified farms when regression models and PSM (propensity-score matching) analysis feature controls for production scale, as well as other confounding factors.

Quite apart from the specific plight of workers employed by smallholder farmers, an important factor is whether the labour standards applied exceed national laws and are properly enforced or not. Qualitative research suggests this is largely dependent on national context and several reviewed studies reported cases in which this assumption does not hold (Moore 2010, Valkila and Nygren 2008, Nelson and Martin 2013, Raynolds 2012, Smith 2010). The problem for standard setters is that this is a highly contextual factor that lies outside their control. However, some CS are striving to go beyond such basic standards and establish a ‘living wage’, which entails basing minimum acceptable wages on an analysis of the living costs faced by the potential beneficiaries of certification.²¹ This is resource- and time-consuming and therefore not all CS may be able to enforce such an initiative.

²⁰ After the review was completed two more recent studies focused on this issue (Risgaard and Okinda 2018, Nelson and Phillips 2017).

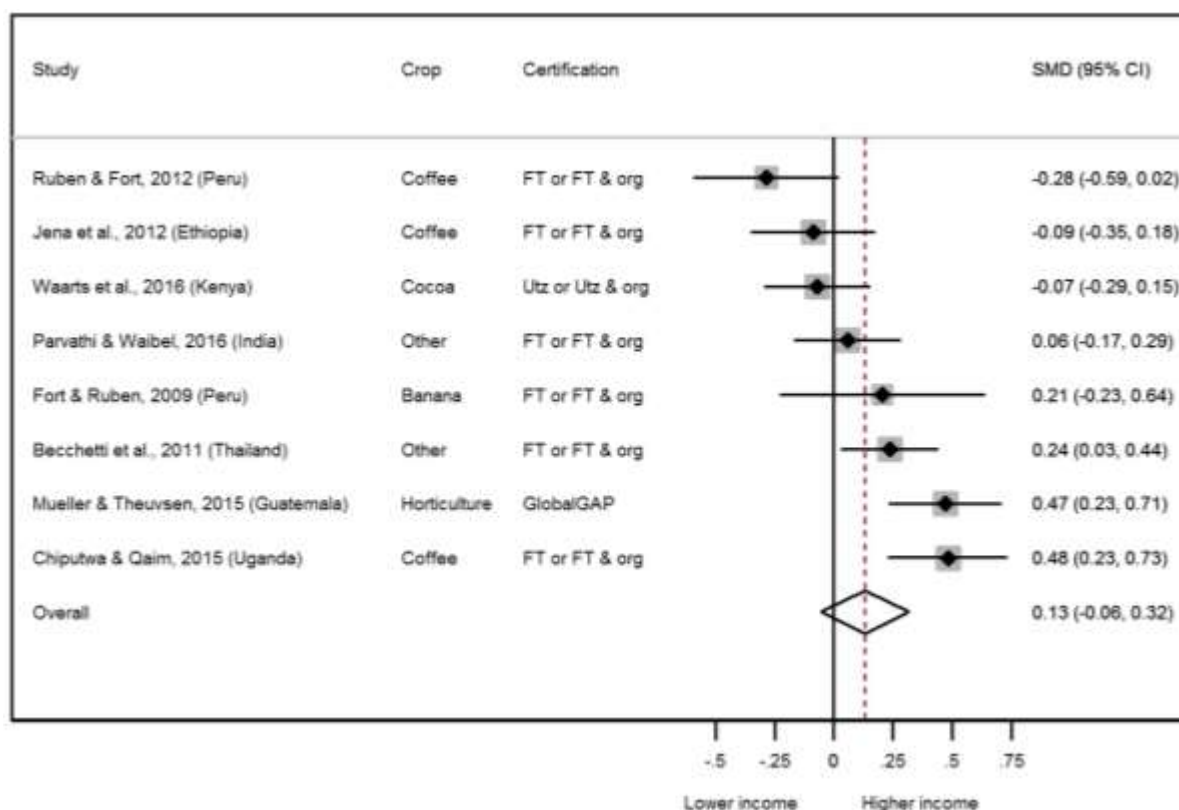
²¹ The best known response in this regard is the Global Living Wage Coalition, which includes six CS that are ISEAL members, which ‘recognise that a living wage is crucial to their individual certification programmes and they have agreed to a shared approach for measuring living wage’. See <http://www.isealalliance.org/our-work/improving-effectiveness/global-living-wage-coalition>

Much of the quantitative and qualitative research reviewed for this study also pays attention to non-wage working conditions such as benefits, health & safety issues and workers' empowerment. In fact, for some CS these aspects of the labour relation are as important as remuneration and wages. With regards to non-wage effects the evidence is mixed. Some studies report worse conditions across a number of labour standards, from health and safety, to collective bargaining and employment security (Cramer et al. 2014a, Luetchford 2008, Moore 2010), whereas others suggest certified farms tend to offer better workplace conditions in terms of benefits and security (Smith 2010; Raynolds 2014, Nelson & Martin 2013). The latter tends to reflect the fact that CS make certified plantations more visible to auditors and pressure groups, precisely because they carry an international certification, and some researchers argue that there are sometimes greater pressures for better labour conditions in certified value chains (Smith 2010, Raynolds 2014).

4.4. Effects on household income

Total household income comprises all sources of income that contribute to the livelihood of a given household, including income from (certified and non-certified) farming, agricultural and non-agricultural employment, and any other non-farm activities. The mixed results on income from certified production and wages suggest the impact on household income may also be uncertain. We found no overall effect on household incomes, as the estimated pooled effect is not statistically distinguishable from zero, though the point estimate is positive. The effect estimates provided by individual studies range from negative to positive, but the only statistically significant point estimates are all positive (see Figure 8). The spread of results for studies examining Fairtrade certification is notable, with SMD point estimates from individual studies ranging from -0.28 (Ruben & Fort 2012) to 0.48 (Chiputwa et al. 2015). The overall null finding for this outcome is very robust to subgroup analysis and separately analysing studies by risk of bias rating, peer review status or level of financial independence from certification bodies produces null findings across all subgroups.

Figure 8 - Forest plot for total household income



Even where effects on income from certified production are positive, the impact on total household income is uncertain and depends on the dynamics of wider farm and labour incomes, and especially on substitution effects between different sources of income. Household income is also dependent on incomes from the sale of non-certified products, especially considering the aforementioned demand constraints affecting certified products. One challenge in understanding the linkages between these different sets of effects is that there is only partial overlap between studies reporting on household income and those that report on income from certified production or wages. Thus we are comparing different sets of studies.

Qualitative evidence suggests two main reasons why effects on certified income do not translate into effects on total household income. One is the degree of reliance of individual households on income from certified production, which may provide only a small share of total income in some cases. Instead, other sources of income, including off-farm employment opportunities, may be more significant and thus weight more on the overall household income. Additionally, other external forms of support, such

as government or NGO aid to producers, often channelled through certified POs (a phenomena described as the ‘honey pot effect’, see Nelson et al. 2002 and Griffiths 2012) may also alter certification effects on the total household income. Second, qualitative evidence suggests that increased costs of certified production, particularly when organic certification is involved, can offset price benefits and moderate the impact on household income. There are several reasons why certified production is more costly for producers. Certified production is more labour intensive as producers are required to meet increasingly strict and demanding standards (Shreck 2002, Smith 2007, Laroche et al. 2012, Moberg 2008). This can increase the use of hired labour (Trauger 2014 Setrini 2011), as well as family labour (Jaffee 2006 Nelson et al. 2002) and even child labour (Dragusanu and Nunn 2014, Valkila and Nygren 2008). Moreover, CS-related paperwork, meetings and trainings, are reported to significantly increase producers' workload (Walsh 2004, Brown 2012, Gomez-Cardona 2012, Pongratz-Chander 2007). The adoption of standards can also lead to product losses when productivity is negatively affected.. Finally, inspection and certification costs can be significant and POs may deduct them from price premium if no external financial support is available (Rueda and Lambin 2013, Dowdall 2012).

4.5.Context matters

The causal chains discussed here confirm that a multiplicity of factors shape impact, from the specific designs of the different interventions, to the fidelity of their implementation and the contextual elements that affect who benefits, when and how. To illustrate how different barriers and contextual factors relate to aspects of implementation, distribution, unintended consequences and a range of other mediating factors we highlight four cross-cutting issues that arise in a large number of context and scheme types. The full range of barriers, enablers and external contextual factors we identified are summarised in Appendix 3.

First, a common feature in almost all CS is uneven participation across economically and socially heterogeneous groups of producers and workers. Poorer and more vulnerable participants, including women, are generally harder to reach and hence less frequently included. This is related to barriers to entry, which exist for all CS, as well as to the systems’ own preferences for producers that are already

organised in established POs. This is particularly the case in CS that are more focused on supply chain performance, on quality and timeliness, than in CS that are more oriented towards supporting vulnerable producers. Working with groups that are already organised facilitates the process of certification and auditing, especially when dealing with large numbers of scattered smallholders, and can lower the administrative burden facing certifying bodies.

Second, the direct costs of obtaining the certification itself are frequently mentioned as an important barrier to entry, although there is huge variation in certification costs across CS. Even when certification costs are relatively modest, they can still exclude the poorest and smallest producers, particularly considering the additional investment costs required to meet standards.

Third, when certification operates through POs, and certification rents are collectively managed (e.g. premiums for social investments), there is little evidence to suggest that CS can avoid or overcome (pre-existing) local conflicts (Naylor 2014, Pollack 2006), a lack of common interests among members (Naylor 2014, Staib 2012, Sen 2009), unequal gender relations (Lyon et al. 2010, Stathers and Gathuthi 2013, Terstappen 2010, TWIN 2013, Stenn 2015) or elite capture (Sutton 2014, Cramer et al. 2014b, Staib 2012, Dolan 2010). The latter can lead to the manipulation and control of POs by members of their own management. CS do not operate in a social, institutional or political vacuum, and radical changes to deep-rooted socio-economic dynamics are seldom triggered by the access to certification.

Fourth, there is a body of evidence that calls into question the effectiveness of monitoring and auditing systems, which are essential for proper targeting and to ensure compliance with standards. The points raised include a lack of transparency and accountability of auditing bodies towards audited producers (Lyon 2005, Setrini 2011, Larsen et al. 2014), the degree to which external audits can be manipulated by POs and plantation management to simulate compliance (Heller 2010, Moore 2010, Schelly 2011), and the lack of the auditor's ability to grasp the producers' everyday reality and make socio-culturally appropriate recommendations (Heller 2010, Trauger 2014, Sen 2009). Shortcomings in these areas can affect both compliance and the ability of monitoring systems to provide appropriately tailored feedback and recommendations to improve practices. It is therefore important to understand the effects of CS in

terms of the realities of their networks and the multiple actors that play a role in ensuring certification is implemented according to agreed protocols but often with limited resources or inadequate incentives (extension agents, auditors, processing intermediaries, NGO workers, etc.).

Finally, context may matter differently for different types of interventions. In some contexts price interventions may be less effective than capacity building interventions and *vice versa*. Estimating the impact of, say, Fairtrade certification, will however not tell us whether the impact of guaranteed prices is more important than sustained capacity building for a PO because the context was more favourable for that kind of intervention. Given that most CS expand the range of standards and interventions in any given setting, it is important to use rigorous mixed methods approaches to try to disentangle the relative effectiveness of particular standards and interventions and the contextual factors that affect each causal chain.

5. Conclusions

This systematic review synthesized evidence of the impact of multiple certifications featuring many different interventions across on a wide range of crops and with two broadly defined groups of participants: agricultural producers and workers in LMICs. The effectiveness review synthesized quantitative effects and found positive impact on prices, incomes from certified production and schooling, which indicate that CS can improve some intermediate outcomes for direct producers. However, further along the causal chain the evidence we found is inconclusive for overall household incomes and we find no evidence of improvement in wages for farm workers. This is a mixed picture that can be interpreted in a number of ways depending on how a wide range of contextual factors that shape the effectiveness of interventions by standards systems (or lack thereof) are accounted for. There is a general consensus that the impact of certification is context specific, but this mixed-method review provides insights into key contextual factors affecting particular causal chains, such as certified farm income, workers' wages and total household income.

One limitation of the evidence base is that the effectiveness results can only be interpreted in terms of whether certifications, as bundles of interventions, have any impact. We would of course prefer to learn which specific individual interventions (such as the price premium or requirements of good agricultural practices) affect key outcomes. The proliferation of different standards and interventions within every certification system (e.g. Fairtrade, Utz and Rainforest Alliance, Global G.A.P. among others) makes it difficult to link a specific standards system to a specific type of intervention. Most schemes combine a mix of capacity building interventions, price and marketing support together with enforced requirements regarding practices of production and employment. The next generation of impact evaluation studies could then attempt to disentangle the effects of specific types of interventions within each standards system.

Despite the exhaustive and systematic search for relevant material published between 1990 and 2015, and despite the large volume of literature on the impact of agricultural certification, the review only found a limited number of eligible quantitative studies for each of the selected socio-economic outcomes, and some CS (in particular Fairtrade) are disproportionately represented in the evidence base. While we included 43 studies in the effectiveness review, these were distributed across several outcomes, so the number of studies per outcome was limited. Two important reasons are the scarcity of rigorous study designs that address selection bias and control for confounding factors, and the generally weak standards in reporting key statistical information that is required to synthesize effect sizes. The number of rigorous quantitative impact evaluations has grown substantially since 2008, which goes some way to addressing the first problem. Solving the second problem will require building a common understanding of reporting standards. By contrast, the number of qualitative studies included in the review is substantial and these provide a wealth of information on barriers and facilitators as well as on other prominent contextual aspects, evidence of which is integrated in the analysis of quantitative effects. However, much of this qualitative evidence is extracted from studies that were not combined with a rigorous quantitative evaluation in the same project or intervention. This calls for more mixed-meth-

ods evaluations, combining quantitative methods that can address issues of causal attribution with qualitative/ethnographic methods that can contribute much-needed evidence on causal pathways, mechanisms and unequal and unintended effects of certification in agriculture.

Despite these limitations, this review has shown that voluntary standards can potentially contribute to positive outcomes and gradually improve the position of agricultural producers and workers. In most settings they require the right conditions to work, and not all certification systems are equally effective. Context matters and shows the limitations of the certification model, since it requires institutional and economic environments conducive to more agricultural dynamism, tighter labour markets through expanded job creation, better coordinated supply chains to reach socio-economic sustainability, and greater demand for certified products. It is difficult to find situations in which all these elements are in place. These factors relate to the broader agricultural and economic policy environment, which is generally outside the control of individual standards systems. Moreover, many certification systems are not merely agents of ‘development interventions’ but rather complex networks involving organisations in high-income and LMICs, with a variety of interests, power relations and institutional frameworks that are endogenous to the nature of and interrelations between their participants, whether individuals or collective organisations (Paschall 2013). The diversity of CS, despite their common social sustainability goals, is also exacerbated by the gradual transition from ‘niche’ to mainstream markets, which entails an ever growing role of key market actors, such as retail chains (Dolan 2010; Lernoud et al. 2017). Their outcomes cannot be analysed in isolation from these factors.

These results also call into question the effectiveness of agricultural certification in helping the poorest of poor. The implication is that CS need to manage expectations and carefully frame their claims about impact, especially when positive effects are only marginal or even non-existent. The difficulty faced by certification systems is the latent tension between the market imperative to promote a particular certification, which may lead to bold claims about sustainability, and the operational imperative to improve the evidence base on impact through more rigorous and more frequent impact evaluations.

References

- Abarca-Orozco, S.J. (2015). *Production and marketing innovations in Fair Trade and organic coffee cooperatives in the Cordoba-Huatusco corridor in Veracruz, Mexico*. PhD. Iowa State University.
- Aidenvironment. (2016). *Evaluation of Utz in the Indonesian cocoa sector*. Amsterdam: Aidenvironment.
- Aidenvironment (2017). The business benefits of using sustainability standards A meta-review. Aidenvironment report for ISEAL Alliance. Retrieved from <http://www.aidenvironment.org/wp-content/uploads/2017/03/Aidenvironment-report-business-benefits-of-standards.pdf>
- Amekawa, Y. (2001). Can a public GAP approach ensure safety and fairness? A comparative study of Q-GAP in Thailand. *Journal of Peasant Studies*, 40(1), pp.189-217.
- Anteneh, A., Muradian, R., & Ruben, R. (2014). The more the better? The effects of multiple certification on smallholder coffee livelihood: evidence from southern Ethiopia. In: Hoebink, P., Ruben, R., Elbers, W., & van Rijsbergen, B., ed. 2014, *The impact of coffee certification on smallholder farmers in Kenya, Uganda and Ethiopia*. Radboud University Nijmegen: Centre for International Development Issues Nijmegen.
- Arce, A. (2009). Living in Times of Solidarity: Fair Trade and the Fractured Life Worlds of Guatemalan Coffee Farmers. *Journal of International Development*, 21,1031-1041.
- Asfaw, S., Mithoefer, D., & Waibel, H. (2008). EU private agrifood standards in African high- value crops: pesticide use and farm-level productivity. *12th Congress of the European Association of Agricultural Economists (EAAE)*, 26-29.
- Asfaw, S., Mithofer, D., and Waibel, H. (2009). Investment in Compliance with GlobalGAP Standards: Does It Pay Off for Small-Scale Producers in Kenya? *Quarterly Journal of International Agriculture*, 48(4), pp.337-362.
- Asfaw, S., Mithoefer, D., & Waibel, H. (2010). Agrifood supply chain, private-sector standards, and farmers' health: evidence from Kenya. *Agricultural Economics*, 41, 251-263.
- Atkins, S., Lewin, S., Smith, H., Engel, M., Fretheim, A., & Volmink, J. (2008). Conducting a meta-ethnography of qualitative literature: lessons learnt. *BMC medical research methodology*, 8(1), 21.
- Auld, G., Renckens, S., & Cashore, B. (2015). Transnational private governance between the logics of empowerment and control. *Regulation & Governance*, 9(2), 108-124.

- Babin, N. (2012). *Agrarian Change, Agroecological Transformation and the Coffee Crisis in Costa Rica*. PhD dissertation. University of California.
- Bacon, C.M. (2005). *Confronting the coffee crisis: Nicaraguan farmers use of cooperative, Fair Trade and agroecological networks to negotiate livelihoods and sustainability*. PhD dissertation. University of California, Santa Cruz.
- Bagama, F., Hoebink, P., Elber, W., van Kempen, L., and van Rijsbergen, B. (2014). The Impact of UTZ Certification on Smallholder Farmers in Uganda. In: Hoebink, P., Ruben, R., Elbers, W., and van Rijsbergen, B., ed. 2014, *The impact of coffee certification on smallholder farmers in Kenya, Uganda and Ethiopia*. Radboud University Nijmegen: Centre for International Development Issues Nijmegen.
- Bakker, P. (2014). *Do Smallholders Benefit from Group Certification? A Case Study on the Social and Economic Effects of Rainforest Alliance Group Certification on Small-Scale Tea Farmers in Misiones, Argentina*. Masters dissertation. Wageningen UR.
- Balineau, G. (2011). Le renforcement des organisations de producteurs de coton au Mali: Enjeux, impact et leçons du commerce équitable pour la privatisation de la filière. *Canadian Journal of Development Studies*, 32(3), 297-312.
- Barham, B.L., & Weber, J.G. (2012). The Economic Sustainability of Certified Coffee: Recent Evidence from Mexico and Peru. *World Development*, 40(6), 1269-1279.
- Barrientos, S., Dolan, C., & Tallontire, A. (2003). A gendered value chain approach to codes of conduct in African horticulture. *World Development*, 31(9), 1511–1526.
- Beall, E. (2012). Smallholders in Global Bioenergy Value Chains and Certification: Evidence from Three Case Studies. *Environment and Natural Resources Management Working Paper*, 50.
- Becchetti, L., & Gianfreda, G. (2008). Does Consumption Heal Producers? The Effect of Fair Trade on Marginalised Producers' Health in Kenya. *Cahiers Agricultures*, 19, 23-27.
- Becchetti, L., Castriota, S., & Michetti, M. (2008). Child labour and consumer responsibility: an impact study. Working Paper 103, Society for the Study of Economic Inequality.
- Becchetti, L., Conzo, P, and Gianfreda, G. (2011). Market access, organic farming and productivity: the determinants of creation of economic value on a sample of Fair Trade affiliated Thai farmers. *The Australian Journal of Agricultural and Resource Economics*, 56, pp.117–140.
- Bennett, M., Francesconi, G.N., Giovannucci, D., & Daitchman, J. (2012). *Cote d'Ivoire Cocoa: A COSA survey of Rainforest Alliance certified farms*. Committee on Sustainability Assessment.
- Bergeron, M. (2010). “Para dar mas titulo a las mujeres”: Le potentiel d'empowerment du commerce équitable de cafe pour les productrices costaricaines. Masters. University of Ottawa.

- Besky, S. (2014). *The Darjeeling distinction: labor and justice on fair-trade tea plantations in India*. Berkeley: University of California Press.
- Beuchelt, T., Zeller, M., & Oberthur, T. (2009). Justified hopes or utopian thinking? The suitability of coffee certification schemes as a business model for small-scale producers. In: International Association of Agricultural Economists Conference, Beijing, China, 16-22 August 2009.
- Blackman, A., & Rivera, J. (2010). *The evidence base for environmental and socioeconomic impacts of 'sustainable' certification*. Washington D.C.: Resources for the Future.
- Bonanno, A., and Cavalcanti, J.S.B. (2012). Globalization, food quality and labor: the case of grape production in North-eastern Brazil. *International Journal of Sociology of Agriculture and Food*, 19(1), pp.37-55.
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., and Rothstein, H. R. (2009). *Introduction to meta-analysis*. Chichester: John Wiley & Sons.
- Brown, S. (2012). One hundred years of labor control: violence, militancy, and the Fairtrade banana commodity chain in Colombia. *Environment and Planning A*, 45(11), pp.2572-2591.
- Byerlee, D., & Rueda, X. (2015). From public to private standards for tropical commodities: a century of global discourse on land governance on the forest frontier. *Forests*, 6, 1301-24.
- Campbell, R., Pound, P., Pope, C., Britten, N., Pill, R., Morgan, M., & Donovan, J. (2003). Evaluating meta-ethnography: a synthesis of qualitative research on lay experiences of diabetes and diabetes care. *Social science & medicine*, 56(4), 671-684.
- Carimentrand, A., and Ballet, J. (2010). When Fair Trade increases unfairness: The case of quinoa from Bolivia. *Cahier FREE*, 5.
- Cepeda, D., Pound, B., Nelson, V., Kajman, G., Cabascango, D., Martin, A., Chile, M., Posthumus, H., Caza, G., Mejia, I., Montenegro, F., Ruup, L., Velastegui, G.A., Tiaguaro, Y., Valverde, M., & Ojeda, A. (2013). *Assessing the Poverty Impact of Sustainability Standards: Ecuadorian Cocoa*. University of Greenwich: Natural Resources Institute.
- Center for Evaluation of the Saarland University (CESU). (2012). *Assessing the Impact of Fairtrade on Poverty Reduction through Rural Development*. Saarland University.
- Chiputwa, B., & Qaim, M. (2014). Sustainability standards, gender, and nutrition among smallholder farmers in Uganda. GlobalFood Discussion Papers, RTG1666, Georg-August- University of Göttingen.
- Chiputwa, B., Spielman, D. J., & Qaim, M. (2015). Food standards, certification, and poverty among coffee farmers in Uganda. *World Development*, 66, 400-412.

- Cofre, G., Riquelme, I., Engler, A., & Jara-Rojas, R. (2012). Adopción de Buenas Prácticas Agrícolas (BPA): costo de cumplimiento y beneficios percibidos entre productores de fruta fresca. *IDESIA*, 30(3), 37-45.
- Colen, L., Maertens, M., & Swinnen, J. (2012). Private standards, trade and poverty: GlobalGAP and horticultural employment in Senegal. *World Economy*, 35(8), 1073-1088.
- Cramer, C., Johnston, D., Oya, C., & Sender, J. (2014a). *Fairtrade Employment and Poverty Reduction in Ethiopia and Uganda*. SOAS: The University of London. www.ftepr.org
- Cramer, C., Johnston, D., Oya, C., & Sender, J. (2014b). Fairtrade cooperatives in Ethiopia and Uganda: uncensored. *Review of African Political Economy*, 41(sup1), S115-S127.
- Dannenberg, P., and Nduru, G.M. (2013). Practices in international value chains: the case of the Kenyan fruit and vegetable chain beyond the exclusion debate. *Tijdschrift voor economische en sociale geografie*, 104(1), pp.41-56.
- Darko, E., A. Lynch & W. Smith (2017). *The impact of Fairtrade: a review of research evidence 2009-2015*. ODI: London. <https://www.odi.org/publications/10891-impact-fairtrade-review-research-evidence-2009-2015>
- de Janvry, A., McIntosh, C., & Sadoulet, E. (2014). Fair Trade and Free Entry: Can a Disequilibrium Market Serve as a Development Tool? University of California at Berkeley.
- Dixon-Woods, M., Agarwal, S., Young, B., Jones, D., & Sutton, A. (2004). *Integrative approaches to qualitative and quantitative evidence*. London: Health Development Agency, 181.
- Dragusanu, R., & Nunn, N. (2014). The Impacts of Fair Trade Certification: Evidence From Coffee Producers in Costa Rica. Harvard Papers.
- Dolan, C.S. (2010). Virtual moralities: the mainstreaming of Fairtrade in Kenyan tea fields. *Geoforum*, 41(1), 33-43.
- Donovan, J., and Poole, N. (2014a). Changing asset endowments and smallholder participation in higher value markets: Evidence from certified coffee producers in Nicaragua. *Food Policy*, 44, pp.1- 13.
- Donovan, J., and Poole, N. (2014b). Partnerships in Fairtrade coffee: a close-up look at how buyers and NGOs build supply capacity in Nicaragua. *Food Chain*, 4(1), pp.34-48.
- Dowdall, C.M. (2012). Small farmer market knowledge and specialty coffee commodity chains in Western Highlands Guatemala. PhD. Florida International University.
- Ehlert, C.R., Mithoefer, D., & Waibel, H. (2014). Worker Welfare on Kenyan Export Vegetable Farms. *Food Policy*, 46, 66-73.

- Ellery, R. (2010). Spilling the beans: Exploring women's empowerment in fair trade coffee production. Masters. University of Guelph.
- Fairtrade Foundation. (2013). Branching Out: Fairtrade in Malawi - Monitoring the impact of Fairtrade on five certified organisations. London: Fairtrade Foundation.
- Fairtrade (2015). Fairtrade Theory of Change Version 2.0. Retrieved from https://www.fairtrade.net/fileadmin/user_upload/content/2009/resources/140112_Theory_of_Change_and_Indicators_Public.pdf
- Fayet, L., and Vermeulen, W.J.V. (2014). Supporting smallholders to access sustainable supply chains: lessons from the Indian cotton supply chain. *Sustainable Development*, 22(5), pp.289-310.
- Fort, R., & Ruben, R. (2009). The impact of Fair Trade on banana producers in northern Peru. Paper prepared for: *International Association of Agricultural Economists Conference*, Beijing, China, August 16-22, 42-55.
- Francesconi, G.N., and Ruben, R. (2014). FairTrade's theory of change: an evaluation based on the cooperative life cycle framework and mixed methods. *Journal Of Development Effectiveness*, 6(3), pp.268-283.
- Fraser, J., Fisher, E., and Arce, A. (2014). Reframing 'Crisis' in Fair Trade Coffee Production: Trajectories of Agrarian Change in Nicaragua. *Journal of Agrarian Change*, 14(1), pp.52-73.
- García, C., García, J., Ochoa, G., Mora, J., & Castellanos, J. (2014). *Impact Evaluation of UTZ Certified Coffee Program in Colombia. (2008-2012)*. Colombia: CRECE.
- Getz, C., and Schreck, A. (2006). What organic and Fair Trade labels do not tell us: towards a place-based understanding of certification. *International Journal of Consumer Studies*, 30(5), pp. 490-501.
- Gómez-Cardona, S. (2012). Tensions of Organic Markets for Colombian Coffee Growers: The Case of Valle del Cauca. *Cuadernos de Desarrollo Rural*, 9(68), 65-85.
- Griffiths, P., 2012. Ethical objections to fair trade. *Journal of Business Ethics*, 105(3), 357-373.
- Hannes K. (2011). Chapter 4: Critical appraisal of qualitative research. In Noyes J, Booth A, Hannes K, Harden A, Harris J, Lewin S, Lockwood C (editors), *Supplementary Guidance for Inclusion of Qualitative Research in Cochrane Systematic Reviews of Interventions*. Version 1. Cochrane Collaboration Qualitative Methods Group. Retrieved from <http://cqrmg.cochrane.org/supplemental-handbook-guidance>
- Hansen, H., & Trifković, N. (2014). Food Standards are good—for middle-class farmers. *World Development*, 56, 226-242.

- Hanson, L., Terstappen, V., Bacon, C.M., Leung, J., Ganem-Cuenca, A., Flores, S.R.D., and Rojas, M.A.M. (2012). Gender, health, and Fairtrade: insights from a research-action programme in Nicaragua. *Development in Practice*, 22(2), pp.164-179.
- Heller, A. (2010). Why organizations matter: Certification experiences of coffee producer groups in Guatemala. PhD. Colorado State University.
- Henson, S., & Jaffee, S. (2008). Understanding developing country strategic responses to the enhancement of food safety standards. *The World Economy*, 31(4), 548-568.
- Herman, A. (2010). Connecting the complex lived worlds of Fairtrade. *Journal of Environmental Policy and Planning*, 12(4), pp.405-422.
- International Trade Centre (ITC). (2011). The Impacts of Private Standards on Producers In Developing Countries. (Literature Review Series on the Impacts of Private Standards; Part II). Geneva: International Trade Centre (ITC).
- International Trade Centre (ITC) (2016). Social and Environmental Standards: Contributing to More Sustainable Value Chains. Geneva: ITC/EUI.
- ISEAL (2016). *Researching standards and certification. A guidance note.* https://www.isealalliance.org/sites/default/files/resource/2018-01/Researchers%27_Guidance_Note_Apr2016.pdf
- Jaffee, D.S. (2006). Producing justice: Fair trade coffee, livelihoods and the environment. PhD. The University of Wisconsin - Madison.
- Jari, B., Snowball, J.D., and Fraser, G.C.G. (2013). Is Fairtrade in commercial farms justifiable? Its impact on commercial and small-scale producers in South Africa. *Agrekon*, 52(4), pp. 66-88.
- Jena, P.R., Chichaibelu, B.B., Stellmacher, T., & Grote, U. (2012). The impact of coffee certification on small-scale producers' livelihoods: a case study from the Jimma Zone, Ethiopia. *Agricultural Economics*, 43(4), 429-440.
- Kamau, M.W., Mose, L., Fort, R., & Ruben, R. (2010). The impact of certification on smallholder coffee farmers in Kenya: the case of the Utz certification program. Paper prepared for: The Joint 3rd African Association of Agricultural Economists (AAAE) and 48th Agricultural Economists Association of South Africa (AEASA) Conference, Cape Town, South Africa, September 19-23.
- Kariuki, I.M. (2014). Transition to certification schemes and implications for market access: Global-GAP perspectives in Kenya. *Agricultural Sciences*, 5(12), pp.1100-1111.
- Köhne, M. (2014). Multi-stakeholder initiative governance as assemblage: Roundtable on Sustainable Palm Oil as a political resource in land conflicts related to oil palm plantations. *Agriculture and Human Values*, 31(3), pp.469-480.

- Kuit, M., Guinée, L., & Van Anh. P. (2016). Estimating the impact of implementation of the 4C entry level standard in Uganda and Vietnam. Kuit Consultancy.
- Laroche, K., Jimenez, R., & Nelson, V. (2012). *Assessing the Impact of Fairtrade for Peruvian Cocoa Farmers*. University of Greenwich: Natural Resources Institute.
- Larsen, R.K., Jiwan, N., Rompas, A., Jenito, J., Osbeck, M., & Tarigan, A. (2014). Towards 'hybrid accountability' in EU biofuels policy? Community grievances and competing water claims in the Central Kalimantan oil palm sector. *Geoforum*, 54, 295-305.
- Lernoud, J., Potts, J., Sampson, G., Garibay, S., Lynch, M., Voora, V., Willer, H., & Wozniak, J. (2017). *The State of Sustainable Markets – Statistics and Emerging Trends 2017*. ITC, Geneva.
- Loconto, A.M., and Simbua, E.F. (2012). Making Room for Smallholder Cooperatives in Tanzanian Tea Production: Can Fairtrade Do That? *Journal of Business Ethics*, 108(4), pp.451-465.
- Lipsey, M., & Wilson, D. (2001). *Practical meta-analysis*. London: Sage Publications.
- Luetchford, P. (2008). *Fair trade and a global commodity: coffee in Costa Rica*. London: Pluto Press.
- Lyll, A. (2014). Assessing the Impacts of Fairtrade on Worker-Defined Forms of Empowerment on Ecuadorian Flower Plantations. Switzerland: Fairtrade International and Max Havelaar Foundation.
- Lyon, S. (2005). Maya coffee farmers and the fair trade commodity chain. PhD. Emory University.
- Lyon, S. (2006). Migratory imaginations: the commodification and contradictions of shade grown coffee. *Social Anthropology*, 14(3), pp.377-390.
- Lyon, S. (2007). Fair Trade Coffee and Human Rights in Guatemala. *Journal of Consumer Policy*, 30(3), pp.241-261.
- Lyon, S., Bezaury, J.A., & Mutersbaugh, T. (2010). Gender equity in Fairtrade-organic coffee producer organizations: cases from Mesoamerica. *Geoforum*, 41(1), 93-103.
- Maertens, M., & Swinnen, J. F. (2009). Trade, standards, and poverty: evidence from Senegal. *World Development*, 37(1), 161-178.
- Makita, R. (2011). A confluence of Fair Trade and organic agriculture in southern India. *Development in Practice*, 21(2), pp.205-217.
- Makita, R. (2012). Fair Trade Certification: The Case of Tea Plantation Workers in India. *Development Policy Review*, 30(1), pp.87-107.

- Melkeraen, S. (2009). Black economic empowerment in the South African wine industry: fair trade, power relations and socio-economic rights. Masters. The Norwegian University Library of Life Sciences.
- Mendez, V.E. (2002). Fair Trade Networks in Two Coffee Cooperatives of Western El Salvador. PhD. Colorado State University.
- Mezzadri, A. (2015). Garment Sweatshop Regimes: The Informalisation of Social Responsibility over Health and Safety Provisions. CDPR, SOAS Working paper, 30.
- Milford, A. (2004). Coffee, co-operatives and competition: the impact of fair trade. CMI Report R 2004 6. Bergen: Christian Michelsen Institute.
- Milford, A.B. (2014). Co-operative or coyote? Producers' choice between intermediary purchasers and Fairtrade and organic co-operatives in Chiapas. *Agriculture and Human Values*, 31(4), 577-591.
- Minten, B., Dereje, M., Engeda, E., & Tamru, S. (2015). Who benefits from the rapidly increasing Voluntary Sustainability Standards? Evidence from Fairtrade and Organic certified coffee in Ethiopia. IFPRI Working Paper 71.
- Moberg, M. (2005). Fair trade and Eastern Caribbean banana farmers: rhetoric and reality in the anti-globalization movement. *Human Organization*, 64(1), 4-15.
- Moberg, M. (2008). Slipping away: banana politics and fair trade in the Eastern Caribbean. New York: Berghahn Books.
- Moore, L.B. (2010). Reading Tea Leaves: The Impact of Mainstreaming Fair Trade. Working Paper 10-106. London School of Economics.
- Moyo, P., and Mugabe, T.C. (2014). An appraisal of the socio-economic impacts and challenges of fair trade among emerging farmers in the Eastern Cape, South Africa. *Journal of Human Ecology*, 47(2), pp.201-207.
- Muradian, R., and Pelupessy, W. (2005). Governing the coffee chain: the role of voluntary regulatory systems. *World Development*, 33(12), pp.2029–2044
- Mueller, A.K., & Theuvsen, L. (2015). Financial literacy and food safety standards in Guatemala: The heterogeneous impact of GlobalGAP on farm income. *Global Food Discussion Papers*, 63. Germany: Georg-August-University of Göttingen.
- Naylor, L.B. (2014). Decolonial autonomies: Fair trade, subsistence and the everyday practice of food sovereignty in the highlands of Chiapas. PhD. University of Oregon.

- Nelson, V. & Martin, A. (2011). *Impact Evaluation of Social and environmental Voluntary Standard Systems (SEVSS): Using theories of change*. University of Greenwich: Natural Resources Institute.
- Nelson, V., & Martin, A. (2013). *Final Technical Report: Assessing the Poverty Impact of Sustainability Standards*. University of Greenwich: Natural Resources Institute.
- Nelson, V., Opoku, K., Martin, A., Bugri, J., & Posthumus, H. (2013). *Assessing the poverty impact of sustainability standards: Fairtrade in Ghanaian cocoa*. University of Greenwich: Natural Resources Institute.
- Nelson V. & D. Phillips (2017). Fairtrade, sustainability standards, and the informal economy: What role for hired labourers?. In V. Nelson (ed.), *Fairtrade Impacts: Lessons from around the world*. Practical Action Publishing.
- Nelson, V. & Pound, B. (2009). *The last ten years: a comprehensive review of the literature on the impact of Fairtrade*. London: Fairtrade Foundation.
- Nelson, V., Tallontire, A., & Collinson, C. (2002). Assessing the benefits of ethical trade schemes for forest dependent people: comparative experience from Peru and Ecuador. *International Forestry Review*, 4(2), 99-109
- Ouma, S. (2010). Global Standards, Local Realities: Private Agrifood Governance and the Restructuring of the Kenyan Horticulture Industry. *Economic Geography*, 86(2), pp.197-222.
- Oya, C., Johnston, D., Evans Muchiri, F., Schaefer, D. S., Dickson, K., & Stansfield, C. (2015). Protocol: Effects of certification systems for agricultural commodity production on socio-economic outcomes in low and middle-income countries: a systematic review. The Campbell Collaboration.
- Oya, C., Schaefer, F., Skolidou, D., McCosker, C., & Langer, L. (2017). Effects of certification schemes for agricultural production on socio-economic outcomes in low-and middle-income countries. *Campbell Systematic Reviews*, 13.
- Paschall, M. (2013). *The Role of Third Party Certification in Improving Small Farmer Livelihoods*. PhD. University of St. Gallen
- Parvathi, P., & Waibel, H. (2016). Organic Agriculture and Fair Trade: A Happy Marriage? Evidence from Black Pepper Farmers in India. *World Development*, 77, 206–220.
- Pawson, R., & Tilley, N. (2004). *Realist Evaluation*. London: British Cabinet Office.
- Pollack, N. (2006). Women's empowerment in Fair Trade coffee co-operatives in Oaxaca, Mexico. Masters. Saint Mary's University.

- Pongratz-Chander, K. (2007). Fair trade coffee in Costa Rica: A new model for sustainable development? PhD. Northern Arizona University.
- Popay, J., Rogers, A., & Williams, G. (1998). Rationale and Standards for the Systematic Review of Qualitative Literature in Health Services Research. *Qualitative Health Research*, 8:3, 341-351.
- Quaadvlieg, J., Roca, M.G., and Ros-Tonen, M.A.F. (2014). Is Amazon nut certification a solution for increased smallholder empowerment in Peruvian Amazonia? *Journal of Rural Studies*, 33, pp.41-55.
- Raynolds, L.T. (2012). Fair Trade Flowers: Global Certification, Environmental Sustainability, and Labor Standards. *Rural Sociology*, 77(4), 493-519.
- Raynolds, L.T. (2014). Fairtrade, certification, and labor: global and local tensions in improving conditions for agricultural workers. *Agriculture and Human Values*, 31(3), 499-511.
- Raynolds, L. T. (2017). Fairtrade labour certification: The contested incorporation of plantations and workers. *Third World Quarterly*, 38(7), 1473-1492.
- Raynolds L.T., and N. Greenfield. (2015). Fair trade: movement and markets. In Raynolds L.T., and E.A. Bennett (ed.) *Handbook of research on fair trade*. Edward Elgar. Pp. 25-41
- Riisgaard, L., Michuki, G., Gibbon, P., & Bolwig, S. (2009). *The performance of voluntary standard schemes from the perspective of small producers in East-Africa*. Copenhagen, Denmark: Danish Institute for International Studies.
- Riisgaard, L., & Okinda, O. (2017). Changing labour power on smallholder tea farms in Kenya. *Competition & Change*, 1024529417742302
- Romanoff, S. (2010). Shade coffee in biological corridors: potential results at the landscape level in El Salvador. *Culture and Agriculture*, 32(1), pp.27-41.
- Ronchi, L. (2002). The Impact of Fair Trade on Producers and Their Organisations: A Case Study with Coocafe in Costa Rica. PRUS Working Paper, 11. University of Sussex.
- Rotter, S. (2000). A just harvest: A naturalistic evaluation of a fair trade coffee project in San Lucas Toliman, Guatemala. Masters. University of Guelph.
- Roy, D, & Thorat, A. (2008). Success in High Value Horticultural Export Markets for the Small Farmers: The Case of Mahagrapes in India. *World Development*, 36(10), 1874-1890.
- Ruben, R. (2012). The Fair Trade Balance: New Challenges after 25 years of Fair Trade. FERDI Policy Brief n. 52. Retrieved from <http://www.ferdi.fr/en/publication/b52-fair-trade-balance-new-challenges- after-25-years-fair-trade>

- Ruben, R. (2013). Critical review of some recently published Fair Trade impact studies: Deficient study design and little robust evidence. Radboud University Nijmegen. Retrieved from: <http://www.ru.nl/cidin/@883941/pagina/>.
- Ruben, R. (2017). Impact assessment of commodity standards: towards inclusive value chains. *Enterprise Development and Microfinance*, 28(1-2), 82-97.
- Ruben, R., Hoebink, P., Ngutu, M.H., Mutwiwa, U., Njuguna, S., Elbers, W., van Kempen, L., & van Rijsbergen, B. (2014). The Impact of Coffee Certification on Smallholder Farmers in Kenya. In: Hoebink, P., Ruben, R., Elbers, W., and van Rijsbergen, B., ed. 2014, *The impact of coffee certification on smallholder farmers in Kenya, Uganda and Ethiopia*. Radboud University Nijmegen: Centre for International Development Issues Nijmegen.
- Ruben, R., & Fort, R. (2012). The impact of fair trade certification for coffee farmers in Peru. *World Development*, 40(3), 570-582.
- Ruben, R., & Zuniga-Arias, G. (2011). How standards compete: comparative impact of coffee certification schemes in Northern Nicaragua. *Supply Chain Management: An International Journal*, 16(2), 98-109.
- Rueda, X., & Lambin, E.F. (2013). Responding to Globalization: Impacts of Certification on Colombian Small-Scale Coffee Growers. *Ecology and Society*, 18(3), 21.
- Ryan, O. (2011). Chocolate nations: living and dying for cocoa in West Africa. London: Zed Books.
- Said-Allsopp, M., and Tallontire, A. (2014). Enhancing Fairtrade for women workers on plantations: insights from Kenyan agriculture. *Food Chain*, 4(1), pp.66-77.
- Sandelowski, M., Docherty, S., & Emden, C. (1997). Focus on qualitative methods. Qualitative meta-synthesis: issues and techniques. *Research in nursing and health*, 20, 365-372.
- Schelly, E. (2011). Representation and partnership: A case study of the worker committee on fair trade certified farms in the Ecuadorian cutflower industry. Masters. Colorado State University.
- Schoonhoven-Speijer, M. (2012). Maintaining a sustainable livelihood; An analysis of the effects of UTZ certification on market access, risk reduction and livelihood strategies of Kenyan coffee farmers. Masters. Radboud University Nijmegen.
- Schuster, M., and Maertens, M. (2013). Do Private Standards Create Exclusive Supply Chains? New Evidence from the Peruvian Asparagus Export Sector. *Food Policy*, 43, pp.291-305.
- Schuster, M., & Maertens, M. (2016). Do private standards benefit workers in horticultural export chains in Peru? *Journal of Cleaner Production*, 112, 2392-2406.

- Schwandt, T.A. (1996) Farewell to criteriology. *Qualitative Inquiry*, 2(1), 58-72.
- Sen, D. (2009). *From illegal to organic: Fair trade-organic tea production and women's political futures in Darjeeling, India*. Rutgers The State University of New Jersey-New Brunswick.
- Setrini, G. (2011). Global niche markets and local development: clientelism and Fairtrade farmer organizations in Paraguay's sugar industry. PhD. Massachusetts Institute of Technology.
- Shreck, A. (2002). Just bananas? A Fair Trade alternative for small-scale producers in the Dominican Republic. PhD. Colorado State University.
- Silva-Castaneda, L. (2012). A forest of evidence: third-party certification and multiple forms of proof - a case study of oil palm plantations in Indonesia. *Agriculture And Human Values*, 29(3), pp.361-370.
- Smith, J. (2007). The Search for Sustainable Markets: The Promise and Failures of Fair Trade. *Culture & Agriculture*, 29(2), 89-99.
- Smith, S. (2010). *Fairtrade Bananas: A Global Assessment of Impact*. University of Sussex: Institute of Development Studies.
- Smith, S., Auret, D., Barrientos, S., Dolan, C., Kleinbooi, K., Njobvu, C., Opondo, M., and Tallontire, A. (2004). *Ethical trade in African horticulture: gender, rights and participation*. University of Sussex: Institute of Development Studies.
- Staib, P.W. (2012). Coffee and the countryside: Small farmers and sustainable development in Las Segovias de Nicaragua. PhD. The University of New Mexico.
- Staricco, J.I., and Ponte, S. (2015). Quality regimes in agro-food industries: a regulation theory reading of Fair Trade wine in Argentina. *Journal of Rural Studies*, (38), pp.65-76.
- Stathers, T., & Gathuthi, C. (2013). *Poverty impact of social and environmental voluntary standard systems in Kenyan tea*. University of Greenwich: Natural Resources Institute.
- Stenn, T.L. (2015). *The Cultural and Political Intersection of Fair Trade and Justice: Managing a Global Industry*. New York: Palgrave Macmillan.
- Sterne, J., Sutton, A., Ioannidis, J., Terrin, N., Jones, D., Lau, J., Carpenter, J., Rücker, G., Harbord, R., Schmid, C., Tetzlaff, J., Deeks, J., Peters, J., Macaskill, P., Schwarzer, G., Duval, S., Altman, G., Moher, D., & Higgins, J. (2011). Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *British Medical Journal*, 343:d4002.
- Subervie, J. & Vagneron, I. (2013). A Drop of Water in the Indian Ocean? The Impact of GlobalGap Certification on Lychee Farmers in Madagascar. *World Development*, 50, 57-73.

- Sutton, S. (2014). Voice, Choice and Governance: The Case of Tanzania's Fairtrade Co-operatives. PhD. Queen Mary, University of London.
- Terstappen, V.L. (2010). A case study of gender, health, and Fair Trade in Nicaragua. Masters. University of Saskatchewan.
- Terstappen, V., Hanson, L., & McLaughlin, D. (2013). Gender, health, labor, and inequities: a review of the fair and alternative trade literature. *Agriculture and Human Values*, 30(1), 21-39.
- Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology*, 8 (45).
- Trauger, A. (2014). Is Bigger Better? The Small Farm Imaginary and Fair Trade Banana Production in the Dominican Republic. *Annals of the Association of American Geographers*, 104(5), 1082- 1100.
- TWIN. (2012). *Making International Supply Chains Work for Smallholder Farmers: A comparative study of six fair trade value chains*. London: Fairtrade Foundation.
- TWIN (2013). *Empowering Women Farmers In Agricultural Value Chains*. London: TWIN.
- Utz (2017). *M&E System. Public M&E System Report Version 3.1*. Retrieved from https://utz.org/wp-content/uploads/2015/12/UTZ_Monitoring-and-Evaluation_System-Report_v3.pdf
- Vagneron, I., & S. Roquigny, (2010). *What do we really know about the impact of fair trade? A synthesis*. Paris: PFCE.
- Valkila, J. (2009). Fair Trade organic coffee production in Nicaragua: Sustainable development or a poverty trap? *Ecological Economics*, 68, 3018-3025.
- Valkila, J., & Nygren, A. (2008). Impacts of Fair Trade-certification on Coffee Farmers, Cooperatives and Laborers, in Nicaragua. *Agriculture and Human Values*, 27(3), 321-333.
- van Rijn, F.C., Judge, L.O., Fort, R., Koster, T., Waarts, Y.R., & Ruben, R. (2016). Fairtrade certification in the hired banana sector. LEI report 2015-056, LEI Wageningen UR.
- van Rijsbergen, B., Elbers, W., Ruben, R., & Njuguna, S.N. (2016). The ambivalent impact of coffee certification on farmers' welfare: a matched panel approach for cooperatives in central Kenya. *World Development*, 77, 277-292.
- Vásquez-León, M. (2010). Free markets and fair trade, collective livelihood struggles, and the cooperative model: two case studies from Paraguay. *Latin American Perspectives*, 37(6), pp.53-73.
- Vogel, I. (2012). Review of the use of 'Theory of Change' in international development. DFID Review Report. London: Department for International Development.

- Waarts, Y., Ge, L., Ton, G., & Jansen, D. (2012). Sustainable tea production in Kenya: impact assessment of Rainforest Alliance and Farmer Field School training. LEI report 2012-043, LEI Wageningen UR.
- Waarts, Y., Lan, G., and Giel, T. (2014). From training to practice: Midterm evaluation of the UTZ-Solidaridad smallholder tea programme in Malawi. LEI report 2013-051. LEI Wageningen UR.
- Waarts, Y., Ingram, V., Liderhof, V., Puister-Jansen, L., & van Rijn, F, and Aryeetey, R. (2016). Impact of UTZ certification on cocoa producers in Ghana, 2011 to 2014. LEI Wageningen UR.
- Walsh, J.M. (2004). Fair Trade in the Fields: Outcomes for Peruvian Coffee Producers. Masters. Massachusetts Institute Of Technology.
- Waddington, H. (2014) Effect size calculation. Presentation in 3ie training programme for systematic reviews. 3ie-International Initiative for Impact Evaluation: London.
- Weber, J.G. (2011). How Much More Do Growers Receive for Fair Trade-Organic Coffee? *Food Policy*, 36(5), 678-685.
- Weiss, C. H. (1997). Theory-based evaluation: Past, present, and future. *New directions for evaluation*, 1997(76), 41-55.
- White, H. (2009). Theory-based impact evaluation: principles and practice. *Journal of development effectiveness*, 1(3), 271-284.
- Wilson, B.R. (2010). Indebted to Fair Trade? Coffee and crisis in Nicaragua. *Geoforum*, 41(1), pp.84-92.

Appendix 1 – Example search strategy: Web of Science-SSCI

[Searched 2 May 2015]

16

1,498

#15 AND #7 AND #4

Indexes=SSCI Timespan=1990-2015

15

370,289

#14 OR #13 OR #12 OR #11 OR #10 OR #9 OR #8

Indexes=SSCI Timespan=1990-2015

14

6,607

TS=((developing or 'less* developed' or 'under developed' or underdeveloped or 'middle income' or 'low* income') NEAR (economy or economies))

Indexes=SSCI Timespan=1990-2015

13

187

TS='transitional countr*'

Indexes=SSCI Timespan=1990-2015

12

2,607

TS=(low NEAR/3 middle NEAR/3 countr*)

Indexes=SSCI Timespan=1990-2015

11

4,039

TS=(lmic or lmics or 'third world' or lamicountr*)

Indexes=SSCI Timespan=1990-2015

10

654

TS=(low NEAR (gdp or gnp or 'gross domestic' or 'gross national' or GNI))

Indexes=SSCI Timespan=1990-2015

9

66,304

TS=((developing or 'less* developed' or 'under developed' or underdeveloped or 'middle income' or 'low* income' or underserved or 'under served' or deprived or poor*) NEAR (countr* or nation? or population? or world or economy or economies))

Indexes=SSCI Timespan=1990-2015

8

326,981

TS=(Afghanistan or Angola or Albania or 'American Samoa' or Argentina or Armenia or Armenian or Azerbaijan or Bangladesh or Belarus or Belize or Benin or Bolivia or Bosnia or Herzegovina or Botswana or Brazil or Bulgaria or Burkina Faso or Burkina Fasso or Burundi or Urundi or Cambodia or Cameroon or Cameroons or Cameron or Camerons or Central African Republic or Chad or Chile or China or Colombia or Comoros or Comoro Islands or Comores or Congo or Costa Rica or Cuba or Zaire or Cote d'Ivoire or Ivory Coast or Djibouti or Dominica* or East Timor or East Timur or Timor Leste or Ecuador or Egypt or United Arab Republic or El Salvador or Eritrea or Ethiopia or Fiji or Gabon or Gambia or Gaza or Georgia Republic or Georgian Republic or Ghana or Grenada or Guatemala or Guinea or Guiana or Guyana or Haiti or Honduras or Hungary or India or Indonesia or Iran or Iraq or

Kazakhstan or Kenya or Kiribati or Korea or Kosovo or Kyrgyzstan or Kirghizia or Kyrgyz Republic or Kirghiz or Kirgizstan or Lao PDR or Laos or Lebanon or Lesotho or Liberia or Libya or Macedonia or Madagascar or Malagasy Republic or Malawi or Malaysia or Maldives or Marshall Islands or Mali or Mauritania or Mauritius or Agalega Islands or Mexico or Micronesia or Moldova or Moldovia or Moldovian or Mongolia or Montenegro or Morocco or Ifni or Mozambique or Myanmar or Myanma or Burma or Namibia or Nepal or Nicaragua or Niger or Nigeria or Pakistan or Palau or Palestine or Panama or Paraguay or Peru or Philippines or Philipines or Phillipines or Phillippines or Romania or Rwanda or Ruanda or Samoa or Samoan Islands or Sao Tome or Senegal or Serbia or Seychelles or Sierra Leone or Sri Lanka or Solomon Islands or Somalia or South Africa or St Lucia or St Vincent or Grenadines or Sudan or Suriname or Swaziland or Syria or Tajikistan or Tadjhikistan or Tadjikistan or Tadjhik or Tanzania or Thailand or Tonga or Togo or Togolese Republic or Tunisia or Turkey or Turkmenistan or Tuvalu or Uganda or Ukraine or Uruguay or Uzbekistan or Uzbek or Vanuatu or Venezuela or New Hebrides or Vietnam or Viet Nam or West Bank or Yemen or Zambia or Zimbabwe)

Indexes=SSCI Timespan=1990-2015

7

210,325

#6 OR #5

Indexes=SSCI Timespan=1990-2015

6

70,400

TS=(coffee OR cocoa OR tea OR infusion* OR 'yerba mate' OR 'camomile' OR sugar* OR fruit* OR banana* OR pineapple* OR mango* OR coconut* OR apricot* OR nut* OR cashew* OR 'shea butter' OR argan OR rice OR quinoa OR bean* OR chickpea* OR 'red kidney' OR lentil* OR soy* OR herb* OR spice* OR 'olive oil' OR olive* OR wine OR honey OR cotton OR flower* OR floriculture OR 'palm oil' OR (crop* NEAR/2 produc*))

Indexes=SSCI Timespan=1990-2015

5

149,299

TS=(Farmer* or farming or agricultur* or horticultur* or grower* or producer* or worker* or labo?rer* or smallholder* or small-holder* or cooperative* or co-operative* or syndicate* or ((trade or labo?r) NEAR union*) or 'agricultural sector' or 'agricultural trade' or 'floriculture' or 'crop production' or 'agricultural products')

Indexes=SSCI Timespan=1990-2015

4

23,229

#3 OR #2 OR #1

Indexes=SSCI Timespan=1990-2015

3

2,207

TS=('fair trade' or fairtrade or fair-trade or transfair or 'fair for life' or 'Rainforest Alliance' or 'Sustainable Agriculture Network' or 'UTZ Certified' or 'UTZ' or 'Global Partnership for Good Agricultural Practice' or 'Global GAP' or 'GlobalGAP' or '4C Association' or 'Better Cotton Initiative' or 'BCI' or 'Cotton made in Africa' or Bonsucro or 'Ethical Tea Partnership' or Trustea or 'International Federation of Organic Agriculture Movements' or IFOAM or 'soil association' or 'IOAS' or 'LEAF' or 'Linking Environment and Farming' or 'Union for Ethical BioTrade' or 'UEBT' or 'Roundtable on Sustainable Palm Oil' or 'RSPO' 'Fair Flowers Fair Plants' or 'ProTerra' or 'ISO 14001')

Indexes=SSCI Timespan=1990-2015

2

16,368

TS=((fair* OR ethic* OR alternative OR sustainab* OR responsib* OR specialty OR eco OR ecologic OR ecological OR organic) NEAR/3 (certifi* OR standard* OR label* OR seal* OR scheme* OR trad* OR market* OR 'value chain*' OR commodit* OR product*))

Indexes=SSCI Timespan=1990-2015

1

6,717

TS=('certification' or 'quality standards' or 'quality label?ing' or 'sustainability standards')

Appendix 2 – Studies included in the quantitative synthesis (RQ1)

| Study | Location | Commodity | Certification scheme(s) | Population | Study design | Method of analysis | Sample size | Overall risk of bias |
|----------------------------|--------------|--|---|------------------------|---|---|-------------------------|----------------------|
| Anteneh et al 2014 | Ethiopia | Coffee | Fairtrade; Organic, UTZ | Agricultural producers | Ex-post controlled observational study | Propensity score matching (PSM) | 700 | High |
| Asfaw et al 2008 | Kenya | Vegetables (French beans, green beans, peas) | EurepG.A.P./ Global-GAP | Agricultural producers | Ex-post controlled observational study | 2SLS/3SLS (Poisson, treatment effects); PSM | 439 | High |
| Balineau 2011 | Mali | Cotton | Fairtrade | Agricultural producers | Retrospective panel built from single data collection exploiting staggered intervention | Fixed effects panel model (GMM estimator) | 198 | High |
| Barham & Weber 2012 | Mexico; Peru | Coffee | Fairtrade; Organic; Rainforest Alliance | Agricultural producers | Panel build from two datasets | Fixed effects panel model | 845(Mexico); 235 (Peru) | Critical |
| Becchetti & Gianfreda 2008 | Kenya | Fruit | Fairtrade; Organic | Agricultural producers | Ex-post controlled observational study | 2 stage treatment model (ordered probit/probit and selection model) | 478 | High |
| Becchetti et al 2008 | Chile | Honey | Fairtrade; Organic | Agricultural producers | Ex-post controlled observational study | GMM regression with LDVs (and a FE model) | 234 | High |
| Becchetti et al 2011 | Thailand | Rice | Fairtrade; Organic | Agricultural producers | Retrospective panel: ex-post controlled observational study with recall questions | 3SLS | 360 | High |
| Cepeda et al 2013 | Ecuador | Cocoa | Rainforest Alliance; Organic; Fairtrade | Agricultural producers | CBA, with baseline and end-line data collection | Difference-in-difference / double difference (DID) | 2010: 576 2012: 415 | Critical |

| Study | Location | Commodity | Certification scheme(s) | Population | Study design | Method of analysis | Sample size | Overall risk of bias |
|----------------------|---------------|-------------------|-------------------------|------------------------|--|---|--|----------------------|
| Chiputwa & Qaim 2014 | Uganda | Coffee | Fairtrade; Organic, UTZ | Agricultural producers | Ex-post controlled observational study | IV model, and simultaneous eqn. model with IV, estimated using maximum likelihood | 419 | Moderate |
| Colen et al 2012 | Senegal | Green bean, mango | GlobalGAP | Wage workers | CBA with baseline and end-line data collection | Cross-sectional OLS; FE panel regression | 163 (for OLS), 46 (for panel) | Moderate |
| Bennett et al 2012 | Cote d'Ivoire | Cocoa | Rainforest Alliance | Agricultural producers | CBA with baseline and end-line data collection | PSM and DID | 2009: 200; 2011: 252 (prior to matching) | Moderate |
| Cramer et al 2014 | Ethiopia | Coffee | Fairtrade | Wage workers | Ex-post controlled observational study | Robust OLS; Propensity score matching (PSM) | 339 | Moderate |
| Cramer et al 2014 | Ethiopia | Flowers | Fairtrade | Wage workers | Ex-post controlled observational study | Robust OLS; Propensity score matching (PSM) | 215 | Moderate |
| Cramer et al 2014 | Uganda | Coffee | Fairtrade | Wage workers | Ex-post controlled observational study | Robust OLS; Propensity score matching (PSM) | 237 | Moderate |
| Cramer et al 2014 | Uganda | Tea | Fairtrade | Wage workers | Ex-post controlled observational study | Robust OLS; Propensity score matching (PSM) | 199 | Moderate |
| Dragusanu 2014 | Costa Rica | Coffee | Fairtrade | Agricultural producers | Ex-post controlled observational study | Linking of mill-level data with individual HH survey data to estimate FT impact | Unclear, but very large (~40,000) | High |
| Ehlert et al. 2014 | Kenya | Vegetables | GlobalGAP | Wage workers | Ex-post controlled observational study | OLS and MIMIC models | 316 | High |
| Fort & Ruben 2009 | Peru | Banana | Fairtrade | Agricultural producers | Ex-post controlled observational study | Propensity score matching (PSM) | 200 | Moderate |

| Study | Location | Commodity | Certification scheme(s) | Population | Study design | Method of analysis | Sample size | Overall risk of bias |
|-------------------------|-----------|--------------|-------------------------|------------------------|--|--|----------------------------------|----------------------|
| García et al 2014 | Colombia | Coffee | UTZ | Agricultural producers | CBA with baseline and end-line data collection | PSM and DID | 857, much smaller for PSM though | Moderate |
| Jena et al 2012 | Ethiopia | Coffee | Fairtrade; Organic | Agricultural producers | Ex-post controlled observational study | Propensity score matching (PSM) | 249 | High |
| Kamau et al 2010 | Kenya | Coffee | UTZ | Agricultural producers | Ex-post controlled observational study | Propensity score matching (PSM) | 262, after matching | High |
| Kuit et al 2016 | Uganda | Coffee | 4C | Agricultural producers | Ex-post controlled observational study | Propensity score matching (PSM) | 250 | High |
| Kuit et al 2016 | Vietnam | Coffee | 4C | Agricultural producers | CBA with baseline and end-line data collection | Propensity score matching (PSM) | 2006-2008: 152; 2013-2015: 249 | High |
| Minten et al 2015 | Ethiopia | Coffee | Various | Agricultural producers | Ex-post controlled observational study | Robust OLS; Propensity score matching (PSM) | 1600 | Moderate |
| Mueller & Theuvsen 2015 | Guatemala | Peas | GlobalGAP | Agricultural producers | Ex-post controlled observational study | Propensity score matching (PSM) | 276 | Moderate |
| Nelson et al 2013 | Ghana | Cocoa | Fairtrade | Agricultural producers | CBA with baseline and end-line data collection | PSM and DID | 2010: 743, 2012: 697 | Critical |
| Parvathi & Waibel 2016 | India | Black pepper | Fairtrade; UTZ; Organic | Agricultural producers | Ex-post controlled observational study | Multinomial endogenous switching model | 300 | High |
| Riisgaard et al 2009 | Uganda | Coffee | Fairtrade | Agricultural producers | Ex-post controlled observational study | 2-stage treatment model with both OLS and FIML | 149 | Moderate |
| Riisgaard et al 2009 | Uganda | Coffee | UTZ | Agricultural producers | Ex-post controlled observational study | 2-stage treatment model with both OLS and FIML | 147 | Moderate |
| Roy & Thorat 2008 | India | Grapes | EurepGAP/ GlobalGAP | Agricultural producers | Ex-post controlled observational study | 2SLS (IV) | 183 | High |

| Study | Location | Commodity | Certification scheme(s) | Population | Study design | Method of analysis | Sample size | Overall risk of bias |
|---------------------------|--------------------|--|--|------------------------|--|---------------------------------|----------------------|----------------------|
| Ruben & Fort 2012 | Peru | Coffee | Fairtrade; Organic | Agricultural producers | Ex-post controlled observational study | Propensity score matching (PSM) | 360 | Moderate |
| Ruben & Zúñiga-Arias 2011 | Nicaragua | Coffee | Fairtrade; Organic | Agricultural producers | Ex-post controlled observational study | Propensity score matching (PSM) | 315 | Unclear |
| Ruben et al 2014 | Kenya | Coffee | Fairtrade; UTZ | Agricultural producers | CBA with baseline and end-line data collection | PSM and DID | 2009: 600; 2013: 493 | High |
| Schuster & Maertens 2014 | Peru | Horticulture (asparagus, grapes, avocado, artichoke) | Various (29 different private standards) | Wage workers | CBA with baseline and end-line data collection | Instrumental variable | 2013: 592; 2014: 499 | Low |
| Stathers & Gathuthi 2013 | Kenya | Tea | Fairtrade, Rainforest Alliance | Agricultural producers | CBA with baseline and end-line data collection | Difference-in-difference | 2010: 897; 2012: 700 | Critical |
| Subervie & Vagneron 2013 | Madagascar | Lychee | GlobalGAP | Agricultural producers | Ex-post controlled observational study | PSM and DID | 505 | High |
| van Rijn 2016 | Ghana | Banana | Fairtrade | Wage workers | Ex-post controlled observational study | Propensity score matching (PSM) | 326 | Critical |
| van Rijn 2016 | Dominican Republic | Banana | Fairtrade | Wage workers | Ex-post controlled observational study | Propensity score matching (PSM) | 258 | High |
| van Rijn 2016 | Colombia | Banana | Fairtrade | Wage workers | Ex-post controlled observational study | Propensity score matching (PSM) | 431 | High |
| Van Rijsbergen et al 2016 | Kenya | Coffee | Fairtrade; UTZ | Agricultural producers | CBA with baseline and end-line data collection | PSM and DID | 218 | Low |
| Waarts et al 2012 | Kenya | Tea | Rainforest Alliance | Agricultural producers | CBA with baseline and end-line data collection | Difference-in-difference (DID) | 331 | High |
| Waarts et al 2016 | Ghana | Cocoa | UTZ | Agricultural producers | CBA with baseline and end-line data collection | PSM and DID | 352 | High |

| Study | Location | Commodity | Certification scheme(s) | Population | Study design | Method of analysis | Sample size | Overall risk of bias |
|------------|----------|-----------|-------------------------|------------------------|--|---|-------------|----------------------|
| Weber 2011 | Mexico | Coffee | Fairtrade; Organic | Agricultural producers | Ex-post controlled observational study | Treatment effects model estimated with maximum likelihood | 845 | Moderate |

Appendix 3 – Summary of findings from qualitative synthesis (RQ2)

| IMPLEMENTATION DYNAMICS | |
|--|---|
| Targeting and (self-) selection of participants | <p>Wealth Capacity of POs, producers and plantations to bear the extra costs of certified production, withstand payment delays and obtain external financial support are key in CS adoption. Production capacity, land size, household size, education and literacy skills, and degree of market integration influence participation in CS. Findings point to systematic pre-existing differences in wealth and resources between certified producers, POs and plantations and non-certified or newcomers.</p> <p>Gender Female participation in CS is undermined by difficulties in combining household work with certification-related activities, lower education and lack of skills, as well as socio-cultural norms. Women-only POs may provide a more protected environment that enhances female participation.</p> |
| Certification interventions and their implementation | <p>Training, new farming practices and PO support Tailoring training to producers' characteristics and needs, and providing sufficient and skilled extension staff are key elements of successful training. A general lack of producers' knowledge and understanding of CS is reported. Cost of applying new practices appears to be main barrier to adoption of GAP. Certified POs tend to receive financial and technical support from a variety of actors, not always related to CS.</p> <p>Pre-payment & credit The pre-payment standard is not always properly enforced. CS appear to lead to better access to credit for POs, produces and workers. Pre-payment and credit are often reported insufficient to cover costs of certified production. CS-related credit more important during crop crises when other financial entities cease support</p> <p>Minimum Price Effectiveness depends on price volatility. During price slumps it allows POs to maintain their market share and provides stability, but becomes irrelevant during price spikes. Oversupply of certified products affects the protection from the minimum price mechanism, as protection is related to the portion of production sold to the certified market. Even when above the conventional price, the minimum price is often reported to be insufficient to cover costs of certified production.</p> <p>Price premium Deductions in the final payment to the producer (PO administrative and certification costs, debt cancellations, price boosters) and oversupply affect the significance of the premium. The collective management of the "social" premium can be undermined by local conflicts, lack of common interests, elite capture, and control and manipulation by plantations' management.</p> <p>Labour standards Effectiveness of CS is related to local labour legislation and its enforcement. Strong legislation may make CS unnecessary, weak legislation may allow CS to bring significant improvements. Overtime restrictions appear to be conflicting with the interests of workers and plantations, particularly when minimum wage is below the living wage.</p> |
| Costs of certified production | <p>Production under organic-social standards significantly increases the use of hired and family labour. Poorer producers may rely more on child labour. CS-related paperwork, meetings and trainings, are also reported to significantly increase producers' workload. Stricter quality criteria and transition from conventional to organic production can cause product losses and therefore increase the cost of production. Inspection and certification costs are reported to be significant and POs may need external financial support to deal with those.</p> |
| M&A | <p>Internal monitoring systems may present opportunities, particularly for women, to acquire new skills and become actively involved in their PO. Auditing bodies are reported to lack transparency and accountability towards producers, to be limited in grasping producers' reality and in making appropriate recommendations.</p> |
| Spillover & unintended CS effects | <p>CS appear to have an upward influence on local crop prices. Spill-over effects are reported on the adoption of GAP by non-certified producers. Increased labour needs of organic-social standards may increase demand in the local agricultural labour market. CS create demand also for more skilled employment, such as auditors, extension staff, trainers etc. CS-related documentation is reported to be used in land disputes or to create "de facto" property.</p> |

| | |
|---|---|
| Multiple certifi- cation | Overlaps between different standards can influence their effectiveness, both positively and negatively. The dual Fairtrade-organic appears the most widespread combination. While organic certification is often required to access Fairtrade markets, particularly for coffee, increased costs of organic certification can be a barrier to Fairtrade adoption. |
| DISTRIBUTIONAL DYNAMICS | |
| Wealth | The term 'small' producer can be problematic, as it does not recognise that POs' membership base can be heterogeneous in terms of farm size. Larger producers may benefit more from CS, as premia are linked to volumes, are more likely to comply with stricter quality criteria and may dominate the PO decision-making. But mutually-beneficial relationships between larger and smaller producers are also reported, as larger producers allow POs to reach the required efficiencies of scale. |
| Gender | Female contributions in certified production tends to remain invisible, as female producers often lack the assets to register as PO members. CS are related to a gendered increase in workload, which affects disproportionately female producers without an even distribution of benefits. Weak female participation is commonly reported in decision-making related to PO management and premium use, as well as in supervisory/management positions in certified plantations. Gender pay discrimination in certified plantations appears to persist despite certification presence. The effectiveness of Joint Boards as mechanism of empowerment for women workers is questioned by several authors. |
| Wage work- ers | Workers hired by small producers tend to remain invisible in CS and receive no, or very few, benefits. In plantations, temporary workers, as well as migrant and racially discriminated workers, are reported to receive less benefits than their co-workers. |
| CONTEXTUAL BARRIERS AND FACILITATORS | |
| Characteristics of POs, producers and plantations | POs context: management, relationship with producer and with buyers POs' strengths and weaknesses directly affect the effectiveness of CS. Cases of PO mismanagement and corruption appear to be recurrent, affecting producers' participation in CS and the resulting benefits. Producers' ability to understand CS and hold accountable their POs is key in CS effectiveness. Transparency in management and transactions, good credit schemes and extension support are key in enhancing members' loyalty and participation. PO size appears to matter, however, in very context-specific ways. Small PO size improves accountability, but is limited in accessing export markets. Large PO size allows better access to export markets, but losses in service quality and alienation between PO management and membership base are a risk. Externally-imposed POs are more vulnerable to corruption and have weaker links with their members, while POs formed on producers' initiative and efforts have more and better quality members' participation. Producers' propensity to collaborate, high self-confidence and low risk-aversion, and higher education influence CS effectiveness. Long-lasting relations, direct and frequent contact and communication between PO and buyer contribute to CS effectiveness. Engaged buyers may skip CS and offer directly benefits tailored to producers through personalised non-certified channels, if CS are not deemed effective. Plantations: management and workers' committees Plantation management commitment to good social and environmental practices, as well as workers' knowledge of their rights and obligations and of CS mechanisms can enhance CS effectiveness. Workers committees, or Joint Bodies, are reported to have limited decision making power and capacity to act, and may be vulnerable to management manipulation. |
| Markets | CS are reported to improve access to export markets Oversupply of certified products is a common challenge, particularly for Fairtrade, as an important part of certified crops end up at the conventional market. Guaranteed markets as part of a certification initiative are suggested to avoid oversupply. Intermediaries continue to play an important role for certified farmers, as they pay directly upon delivery, can offer advanced payments, purchase at the farm gate and have lower quality criteria. |
| Contexts | Local institutions and politics may hinder or enhance PO formation and performance and hence CS effectiveness. Market liberalisation, disintegration of regulating agencies, and lack of national quality standards leave a void that CS can fill to the benefit of producers. On the contrary, democratic structures, ability to hold governments accountable and strong state regulation policies may limit the role that CS can play. Local power imbalances can affect CS effectiveness and point to tailoring standards to the local context instead of applying a 'one-size-fits-all rule'. CS benefits appear to be more valued in contexts of increased social insecurity and violence. |

Appendix 4 – Studies included in the qualitative synthesis

| Study | Associated papers | Certification scheme(s) | Evidence on implementation dynamics | Evidence on distributional dynamics | Evidence on causal mechanisms and barriers and facilitators | Methodology |
|---------------------|-------------------|-------------------------|--|---|---|--|
| Abarca-Orozco 2015 | n/a | Fairtrade | Material related to: training and new practices; financial premium use | Material related to: uneven adoption of standard | Material related to: cooperative management/performance; governmental legislation, incentives, and/or regulation; markets; production cost for certified goods; | Non-ethnographic. Semi-structured interviews; participant observation. |
| Aidenvironment 2016 | n/a | UTZ | Material related to: training and new practices | None | Material related to: participants' reception/motivation; markets; production cost for certified goods | Non-ethnographic. Mixed methods. |
| Amekawa 2001 | n/a | Q-GAP (Global-GAP) | None | None | Material related to: effective/non-effective adoption; markets | Non-ethnographic. Structured interviews |
| Arce 2009 | n/a | Fairtrade, organic | Material related to: training and new practices | Material related to: distribution of benefits & investments | Material related to: production cost for certified products | Ethnographic. |
| Asfaw et al 2009 | n/a | GlobalGAP | Material related to: implementation costs of certification programmes | None | Material related to: cooperative management/performance; effective/non-effective adoption | Non-ethnographic. Questionnaires, interviews, group discussions. |
| Babin 2012 | n/a | Fairtrade | None | None | Material related to: cooperative management/performance; participants' reception/motivation | Ethnographic. |
| Bacon 2005 | n/a | Fairtrade, organic | Material related to: training and new practices | Material related to: uneven adoption of standard | Material related to: cooperative management/performance; markets; | Non-ethnographic. Surveys. |
| Bagama et al 2014 | n/a | UTZ | Material related to: training and new practices | None | Material related to: cooperative management/performance; participants' reception/motivation | Non-ethnographic. Survey, focus group discussions, interviews. |
| Bakker 2014 | n/a | Rainforest Alliance | Material related to: participant's targeting and self-selection | Material related to: uneven adoption of standard | Material related to: participants' reception/motivation; markets, production cost for certified goods | Non-ethnographic. Interviews, observation, some secondary documentation research |

| | | | | | | |
|--------------------------------|-----|---------------------|---|---|--|--|
| Balineau 2011 | n/a | Fairtrade | Material related to: participant's targeting and self-selection | None | Material related to: participants' reception/motivation | Non-ethnographic. Survey. |
| Barham and Weber 2012 (Mexico) | n/a | Rainforest Alliance | Material related to: training and new practices | None | None | Non-ethnographic. Survey data, cooperative records. |
| Barham and Weber 2012 (Peru) | n/a | Rainforest Alliance | Material related to: training and new practices | None | None | Non-ethnographic. Mixed methods: Survey data, cooperative records. |
| Beall 2012 | n/a | RSPO | Material related to: training and new practices; participant's targeting and self-selection; implementation costs of certification programmes | None | Material related to: participants' reception/motivation; social infrastructure; demand for certified products | Non-ethnographic. "Country level assessments" and semi-structured interviews. |
| Bergeron 2010 | n/a | Fairtrade | None | Material related to: uneven adoption of standard | Material related to: participants' reception/motivation | Non-ethnographic. Semi-structured interviews. |
| Besky 2014 | n/a | Fairtrade | Material related to: certification related services; financial premium use | Material related to: distribution of benefits & investments | Material related to: participants' reception/motivation; governmental legislation, incentives, and/or regulation | Ethnographic. |
| Beuchelt et al 2009 | n/a | Fairtrade, organic | Material related to: financial premium use | None | Material related to: cooperative management/performance; markets; production cost for certified goods | Non-ethnographic. Semi-structured interviews. |
| Bonanno and Cavalcaniti 2012 | n/a | GlobalGAP | Material related to: training and new practices | None | Material related to: governmental legislation, incentives, and/or regulation | Non-ethnographic. Case study methodology- observations, interviews, documents review |
| Brown 2012 | n/a | Fairtrade | Material related to: training and new practices; implementation costs of certification programmes; social premium use | Material related to: distribution of benefits & investments | Material related to: markets | Non-ethnographic. Interviews. |
| Carimentrand and Ballet 2010 | n/a | Fairtrade, organic | None | Material related to: distribution of benefits & investments | None | Non-ethnographic. Interviews. |
| Cepeda et al 2013 | n/a | Fairtrade, organic | Material related to: training and new practices; certification related services; financial premium use; social premium use | Material related to: distribution of benefits & investments | Material related to: cooperative management/performance; participants' reception/motivation; social infrastructure; markets; production cost for certified goods | Non-ethnographic. Survey, questionnaires, interviews, focus groups. |

| | | | | | | |
|---------------------------------------|---------------------|-------------------------|--|---|--|---|
| CESU 2012 (Ecuador) | n/a | Fairtrade | Material related to: training and new practices; certification related services; social premium use | Material related to: distribution of benefits & investments | None | Non-ethnographic. Survey, observation, interviews, focus groups. |
| CESU 2012 (Kenya) | n/a | Fairtrade | Material related to: training and new practices; certification related services; social premium use | Material related to: distribution of benefits & investments | Material related to: social infrastructure | Non-ethnographic. Survey, observation, interviews, focus groups. |
| CESU 2012 (Ghana) | n/a | Fairtrade | Material related to: training and new practices; certification related services; financial premium use; social premium use | Material related to: distribution of benefits & investments | None | Non-ethnographic. Survey, observation, interviews, focus groups. |
| CESU 2012 (Peru; Cofee) | n/a | Fairtrade | Material related to: training and new practices; certification related services; financial premium use; social premium use | Material related to: elite capture | Material related to: cooperative management/performance; social infrastructure | Non-ethnographic. Survey, observation, interviews, focus groups. |
| CESU 2012 (Peru; Banana) | n/a | Fairtrade | Material related to: training and new practices; social premium use | Material related to: distribution of benefits & investments | Material related to: participants' reception/motivation; social infrastructure | Non-ethnographic. Survey, observation, interviews, focus groups. |
| CESU 2012 (India) | n/a | Fairtrade | Material related to: training and new practices; social premium use | Material related to: distribution of benefits & investments | None | Non-ethnographic. Survey, observation, interviews, focus groups. |
| Chiputwa et al 2015 | n/a | Fairtrade, organic, UTZ | Material related to: participant's targeting and self-selection | None | Material related to: cooperative management/performance; markets | Non-ethnographic. Interviews with structured questionnaire. |
| Cofre et al 2012 | n/a | GlobalGAP | Material related to: participant's targeting and self-selection; monitoring and auditing; implementation costs of certification programmes | Material related to: uneven adoption of standard | None | Non-ethnographic. Questionnaire-based sample survey. |
| Cramer et al 2014a (Ethiopia; Coffee) | Cramer et al. 2014b | Fairtrade | None | Material related to: distribution of benefits & investments | Material related to: cooperative management/performance | Non-ethnographic. Mixed methods: surveys, questionnaires, life's work interviews. |

| | | | | | | |
|--|---------------------|--------------------|---|--|---|---|
| Cramer et al 2014a (Ethiopia; Flowers) | Cramer et al. 2014b | Fairtrade | None | Material related to: distribution of benefits & investments | None | Non-ethnographic. Mixed methods: surveys, questionnaires, life's work interviews. |
| Cramer et al 2014a (Uganda; Coffee) | Cramer et al. 2014b | Fairtrade | None | Material related to: distribution of benefits & investments | Material related to: cooperative management/performance | Non-ethnographic. Mixed methods: surveys, questionnaires, life's work interviews. |
| Cramer et al 2014a (Uganda; Tea) | Cramer et al. 2014b | Fairtrade | None | Material related to: distribution of benefits & investments | None | Non-ethnographic. Mixed methods: surveys, questionnaires, life's work interviews. |
| Cramer et al 2014b (Uganda) | Cramer et al. 2014a | Fairtrade | Material related to: certification related services; social premium use | Material related to: distribution of benefits & investments; elite capture | Material related to: effective/non-effective adoption | Non-ethnographic. Mixed methods: market data analysis and interviews. |
| Cramer et al 2014b (Ethiopia) | Cramer et al. 2014a | Fairtrade | Material related to: certification related services; social premium use | Material related to: distribution of benefits & investments | Material related to: effective/non-effective adoption | Non-ethnographic. Mixed methods: market data analysis and interviews. |
| Dannenberg and Nduru 2013 | n/a | GlobalGAP | Material related to: participant's targeting and self-selection; implementation costs of certification programmes | None | Material related to: effective/non-effective adoption | Non-ethnographic. Mixed methods: quantitative survey and interviews. |
| Dolan 2010 | n/a | Fairtrade | Material related to: monitoring and auditing; implementation costs of certification programmes; use of financial premiums; use of social premiums | Material related to: elite capture | Material related to: cooperative management/performance; governmental legislation, incentives, and/or regulation; markets | Non-ethnographic. |
| Donovan and Poole 2014a | Donovan 2014b | Fairtrade, organic | Material related to: training and new practices; certification related services | Material related to: uneven adoption of standard | Material related to: cooperative management/performance; markets | Non-ethnographic. Survey, interviews. |
| Donovan and Poole 2014b | Donovan 2014a | Fairtrade, organic | Material related to: training and new practices; certification related services | None | Material related to: cooperative management/performance; markets | Non-ethnographic. Key informant interviews, household surveys, secondary information. |

| | | | | | | |
|------------------------------|-----|--------------------|--|--|--|---|
| Dowdall 2012 | n/a | Fairtrade, organic | Material related to: training and new practices; certification related services; implementation costs of certification programmes; financial premium use; social premium use | Material related to: distribution of benefits & investments; uneven adoption of standard | Material related to: cooperative management/performance; markets; market volatility; production cost for certified goods | Ethnographic. |
| Dragusanu and Nunn 2014 | n/a | Fairtrade | Material related to: training and new practices; implementation costs of certification programmes | Material related to: distribution of benefits & investments; uneven adoption of standard | None | Non-ethnographic. Secondary archival data. |
| Ellery 2010 | n/a | Fairtrade, organic | Material related to: financial premium use | Material related to: uneven adoption of standard | Material related to: cooperative management/performance | Non-ethnographic. Semi-structured interviews. |
| Fairtrade 2013 (Tea) | n/a | Fairtrade | Material related to: implementation costs of certification programmes; social premium use | Material related to: distribution of benefits & investments | Material related to: cooperative management/performance; social infrastructure; markets; production cost for certified goods | Non-ethnographic. Participatory livelihoods approach, focus group discussions, SSI. |
| Fairtrade 2013 (Ground-nuts) | n/a | Fairtrade | Material related to: training and new practices; certification related services; implementation costs of certification programmes | Material related to: distribution of benefits & investments | Material related to: cooperative management/performance; social infrastructure; markets; production cost for certified goods | Non-ethnographic. Participatory livelihoods approach, focus group discussions, SSI. |
| Fairtrade 2013 (Sugar) | n/a | Fairtrade | Material related to: training and new practices; implementation costs of certification programmes | Material related to: distribution of benefits & investments | Material related to: cooperative management/performance; governmental legislation, incentives, and/or regulation; markets; production cost for certified goods | Non-ethnographic. Participatory livelihoods approach, focus group discussions, SSI. |
| Fairtrade 2015 (Bananas) | n/a | Fairtrade | Material related to: participant's targeting and self-selection | Material related to: distribution of benefits & investments; elite capture | Material related to: cooperative management/performance; participants' reception/motivation | Non-ethnographic. Focus group discussions, SSI, LH. |
| Fairtrade 2015 (Cotton) | n/a | Fairtrade | Material related to: implementation costs of certification programmes; social premium use | Material related to: distribution of benefits & investments; elite capture | Material related to: cooperative management/performance; participants' reception/motivation | Non-ethnographic. Focus group discussions, SSI, LH. |

| | | | | | | |
|----------------------------|--------------|---|--|--|--|---|
| Fairtrade 2015 (Tea) | n/a | Fairtrade | Material related to: implementation costs of certification programmes; social premium use | Material related to: distribution of benefits & investments; elite capture | Material related to: cooperative management/performance | Non-ethnographic. Focus group discussions, SSI, LH. |
| Fayet and Vermeulen 2014 | n/a | Fairtrade, organic, Better Cotton Initiative, Shop for Change | Material related to: implementation costs of certification programmes | None | Material related to: markets | Non-ethnographic. Interviews. |
| Francesconi and Ruben 2014 | n/a | Fairtrade | Material related to: Participant's targeting and self-selection; certification related services; monitoring and auditing; implementation costs of certification programmes; social premium use | Material related to: distribution of benefits & investments | Material related to: cooperative management/performance; governmental legislation, incentives, and/or regulation; markets | Non-ethnographic. Quantitative survey. |
| Fraser et al 2014 | n/a | Fairtrade | Material related to: social premium use | Material related to: elite capture; uneven adoption of standard | Material related to: cooperative management/performance; participants' reception/motivation | Ethnographic. |
| Garcia et al 2014 | n/a | UTZ | Material related to: training and new practices; implementation costs of certification programmes | None | Material related to: cooperative management/performance; participants' reception/motivation; demand for certification services | Non-ethnographic. |
| Getz and Schreck 2006 | Schreck 2002 | Fairtrade | Material related to: training and new practices | Material related to: distribution of benefits & investments; uneven adoption of standard | Material related to: cooperative management/performance; participants' reception/motivation; markets | Non-ethnographic. Questionnaire. |
| Gómez-Cardona 2012 | n/a | Fairtrade, organic | Material related to: training and new practices; monitoring and auditing; implementation costs of certification programmes | None | Material related to: markets | Ethnographic. |
| Hanson et al 2012 | n/a | Fairtrade | None | Material related to: uneven adoption of standard | Material related to: cooperative management/performance | Non-ethnographic. Participant observation, semi-structured interviews, gender-disaggregated focus groups. |

| | | | | | | |
|--------------------|-----|--|---|---|--|---|
| Heller 2010 | n/a | Fairtrade, organic, UTZ, C.A.F.E Practices | Material related to: certification related services; monitoring and auditing; implementation costs of certification programmes; financial premium use; social premium use | Material related to: distribution of benefits & investments | Material related to: participants' reception/motivation; governmental legislation, incentives, and/or regulation; markets; demand for certified products | Ethnographic. |
| Herman 2010 | n/a | Fairtrade | Material related to: social premium use | Material related to: distribution of benefits & investments | Material related to: social infrastructure | Non-ethnographic. Interviews. |
| Jaffee 2006 | n/a | Fairtrade, organic | Material related to: training and new practices; certification related services; implementation costs of certification programmes | None | Material related to: cooperative management/performance; markets | Ethnographic. |
| Jari et al 2013 | n/a | Fairtrade | Material related to: participant's targeting and self-selection; financial premium use; social premium use | Material related to: distribution of benefits & investments | Material related to: cooperative management/performance; markets | Non-ethnographic. Interviews. |
| Jena et al 2012 | n/a | Fairtrade, organic | Material related to: training and new practices; certification related services; social premium use | None | Material related to: cooperative management/performance; participants' reception/motivation | Non-ethnographic. Interviews, survey, focus group discussions. |
| Kariuki 2014 | n/a | GlobalGAP | None | None | Material related to: effective/non-effective adoption | Non-ethnographic. Interviews and questionnaires. |
| Köhne 2014 | n/a | RSPO | Material related to: monitoring and auditing | None | Material related to: governmental legislation, incentives, and/or regulation | Non-ethnographic. Interviews, participatory observation, archival research. |
| Laroche et al 2012 | n/a | Fairtrade | Material related to: certification related services; monitoring and auditing; implementation costs of certification programmes; financial premium use; social premium use | Material related to: uneven adoption of standard | Material related to: cooperative management/performance | Non-ethnographic. Focus groups. |
| Larsen et al 2014 | n/a | RSPO | Material related to: monitoring and auditing | None | None | Non-ethnographic. Qualitative case-study methodology inspired by the |

| | | | | | | |
|-------------------------|---------------------------------------|--------------------|--|--|---|--|
| | | | | | | principles of participatory action research. |
| Loconto and Simbua 2010 | n/a | Fairtrade, organic | Material related to: participant's targeting and self-selection; social premium use | Material related to: distribution of benefits & investments | Material related to: cooperative management/performance; markets; demand for certified products | Non-ethnographic. SSIs, focus group discussions. |
| Luetchford 2008 | n/a | Fairtrade | None | None | Material related to: participants' reception/motivation; markets | Ethnographic. |
| Lyll 2014 | n/a | Fairtrade | Material related to: training and new practice; implementation costs of certification programmes; social premium use | None | None | Non-ethnographic. Focus groups. |
| Lyon 2005 | Lyon 2006, Lyon 2007, Lyon et al 2010 | Fairtrade, organic | Material related to: training and new practices; monitoring and auditing | None | Material related to: demand for certified products | Ethnographic. |
| Lyon 2006 | Lyon 2005, Lyon 2007, Lyon et al 2010 | Fairtrade, organic | Material related to: training and new practices | Material related to: uneven adoption of standard | None | Ethnographic. |
| Lyon 2007 | Lyon 2005, Lyon 2006, Lyon et al 2010 | Fairtrade, organic | Material related to: certification related services | None | None | Ethnographic. |
| Lyon et al 2010 | Lyon 2005, Lyon 2006, Lyon 2007 | Fairtrade, organic | None | Material related to: distribution of benefits & investments; uneven adoption of standard | Material related to: participants' reception/motivation; production cost for certified goods | Non-ethnographic. |
| Makita 2011 | n/a | Fairtrade | None | Material related to: uneven adoption of standard | Material related to: markets | Non-ethnographic. Semi-structured interviews. |
| Makita 2012 | n/a | Fairtrade | Material related to: social premium use | None | Material related to: cooperative management/performance | Non-ethnographic. Semi-structured interviews. |
| Melkeraen 2009 | n/a | Fairtrade | Material related to: implementation costs of certification programmes; social premium use | Material related to: distribution of benefits & investments; elite capture | Material related to: cooperative management/performance | Non-ethnographic. Interviews. |

| | | | | | | |
|----------------------|-------------|---------------------------------|--|--|---|---|
| Mendez 2002 | n/a | Fairtrade, organic | Material related to: certification related services; financial premium use; social premium use | Material related to: elite capture; uneven adoption of standard | Material related to: cooperative management/performance; markets | Non-ethnographic. Focus groups. |
| Milford 2004 | n/a | Fairtrade | Material related to: training and new practices; participant's targeting and self-selection; certification related services; financial premium use | Material related to: uneven adoption of standard | Material related to: cooperative management/performance; participants' reception/motivation; markets; production cost for certified goods | Non-ethnographic. Interviews. |
| Milford 2014 | n/a | Fairtrade, organic | Material related to: participant's targeting and self-selection; certification related services | None | Material related to: cooperative management/performance; participants' reception/motivation; social infrastructure; markets, market volatility | Non-ethnographic. Interviews. |
| Minten et al 2015 | n/a | Fairtrade, Organic, Various VSS | Material related to: participant's targeting and self-selection; financial premium use | None | Material related to: markets; demand for certified products | Non-ethnographic. Surveys and secondary databases. |
| Moberg 2008 | Moberg 2005 | Fairtrade | Material related to: participant's targeting and self-selection; implementation costs of certification programmes; financial premium use; social premium use | Material related to: distribution of benefits & investments; uneven adoption of standard | Material related to: cooperative management/performance; participants' reception/motivation; demand for certified products; production cost for certified goods | Ethnographic. |
| Moberg 2005 | Moberg 2008 | Fairtrade | Material related to: social premium use | Material related to: uneven adoption of standard | Material related to: cooperative management/performance; participants' reception/motivation; markets | Ethnographic. |
| Moore 2010 | n/a | Fairtrade | Material related to: monitoring and auditing; social premium use | None | Material related to: cooperative management/performance; markets | Non-ethnographic. Participant observation and semi-structured interviews. |
| Moyo and Mugabe 2014 | n/a | Fairtrade | Material related to: implementation costs of certification programmes; social premium use | None | Material related to: markets; production cost for certified goods | Non-ethnographic. Interviews and focus group discussions. |
| Naylor 2014 | n/a | Fairtrade, organic | Material related to: monitoring and auditing; financial premium use; social premium use | None | Material related to: cooperative management/performance; social infrastructure; markets; production cost for certified goods | Ethnographic. |

| | | | | | | |
|--|-----|--------------------------------|--|--|---|---|
| Nelson et al 2002 | n/a | Fairtrade | Material related to: training and new practices; certification related services | Material related to: distribution of benefits & investments | Material related to: cooperative management/performance; markets | Non-ethnographic. Interviews. |
| Nelson and Martin 2013 | n/a | Fairtrade | Material related to: certification related services; social premium use | Material related to: distribution of benefits & investments | Material related to: cooperative management/performance | Non-ethnographic. Survey. |
| Nelson and Martin 2013 (Kenya; Wage workers) | n/a | Fairtrade, Rainforest Alliance | Material related to: training and new practices; costs of certification programmes | Material related to: distribution of benefits & investments | Material related to: markets | Non-ethnographic. Interviews. |
| Nelson and Martin 2013 (Kenya; Producers) | n/a | Fairtrade, Rainforest Alliance | Material related to: costs of certification programmes; financial premium use | Material related to: distribution of benefits & investments; elite capture | None | Non-ethnographic. Interviews. |
| Nelson and Martin 2013 (India) | n/a | Fairtrade, Rainforest Alliance | Material related to: certification related services; social premium use | Material related to: distribution of benefits & investments | None | Non-ethnographic. Interviews. |
| Nelson and Martin 2013 (Ecuador) | n/a | Fairtrade, Rainforest Alliance | None | Material related to: distribution of benefits & investments | Material related to: cooperative management/performance | Non-ethnographic. Interviews. |
| Nelson and Martin 2013 (Ghana) | n/a | Fairtrade, Rainforest Alliance | None | Material related to: distribution of benefits & investments | Material related to: cooperative management/performance | Non-ethnographic. Interviews. |
| Ouma 2010 | n/a | GlobalGAP | Material related to: training and new practices; implementation costs of certification programmes; financial premium use | None | Material related to: markets | Non-ethnographic. Field research, interviews, questionnaires, and literature reviews. |
| Pollack 2006 | n/a | Fairtrade, organic | None | Material related to: distribution of benefits & investments; uneven adoption of standard | Material related to: cooperative management/performance; governmental legislation, incentives, and/or regulation | Ethnographic. |
| Pongratz-Chander 2007 | n/a | Fairtrade, organic | Material related to: financial premium use | Material related to: distribution of benefits & investments; uneven adoption of standard | Material related to: cooperative management/performance; governmental legislation, incentives, and/or regulation; markets | Non-ethnographic. Observation. |

| | | | | | | |
|---------------------------------------|---------------|-------------------------|--|---|--|--|
| Quaedvlieg et al 2014 | n/a | Fairtrade, organic | Material related to: training and new practices; implementation costs of certification programmes | None | Material related to: participants' reception/motivation; markets | Non-ethnographic. Semi-structured interviews. |
| Raynolds 2012 | Raynolds 2014 | Fairtrade | Material related to: training and new practices; certification related services implementation costs of certification programmes; social premium use | Material related to: uneven adoption of standard | Material related to: markets | Non-ethnographic. Interviews. |
| Raynolds 2014 | Raynolds 2012 | Fairtrade | None | None | Material related to: participants' reception/motivation | Non-ethnographic. Interviews. |
| Riisgaard et al 2009 (Uganda, Coffee) | n/a | Fairtrade, organic, UTZ | Material related to: certification related services; financial premium use | Material related to: distribution of benefits & investments | Material related to: participants' reception/motivation | Non-ethnographic. Focus group discussions and quantitative household survey. |
| Riisgaard et al 2009 (Uganda, Tea) | n/a | Fairtrade | Material related to: certification related services; financial premium use | Material related to: distribution of benefits & investments | Material related to: participants' reception/motivation | Non-ethnographic. Focus group discussions and quantitative household survey. |
| Riisgaard et al 2009 (Kenya, Tea) | n/a | Fairtrade | Material related to: financial premium use | Material related to: distribution of benefits & investments | Material related to: participants' reception/motivation | Non-ethnographic. Focus group discussions and quantitative household survey. |
| Riisgaard et al 2009 (Kenya; Coffee) | n/a | UTZ | Material related to: certification related services; financial premium use | Material related to: distribution of benefits & investments | Material related to: participants' reception/motivation; production cost for certified goods | Non-ethnographic. Focus group discussions and quantitative household survey. |
| Riisgaard et al 2009 (Ethiopia) | n/a | Rainforest Alliance | Material related to: certification related services; financial premium use | Material related to: distribution of benefits & investments | Material related to: participants' reception/motivation; production cost for certified goods | Non-ethnographic. Focus group discussions and quantitative household survey. |
| van Rijn et al 2016 (Ghana) | n/a | Fairtrade | Material related to: certification related services; financial premium use | Material related to: distribution of benefits & investments | Material related to: social infrastructure | Non-ethnographic. Interviews, survey, gaming sessions. |

| | | | | | | |
|--|-----|---------------------|--|--|---|--|
| van Rijn et al 2016 (Colombia) | n/a | Fairtrade | Material related to: certification related services; financial premium use | Material related to: distribution of benefits & investments | None | Non-ethnographic. Interviews, survey, gaming sessions. |
| van Rijn et al 2016 (Dominican Republic) | n/a | Fairtrade | Material related to: certification related services; financial premium use | Material related to: distribution of benefits & investments | Material related to: social infrastructure | Non-ethnographic. Interviews, survey, gaming sessions. |
| Romanoff 2010 | n/a | Rainforest Alliance | Material related to: training and new practices; implementation costs of certification programmes; financial premium use; social premium use | None | None | Non-ethnographic. Interviews. |
| Ronchi 2002 | n/a | Fairtrade, organic | Material related to: training and new practices; certification related services; financial premium use | Material related to: uneven adoption of standard | Material related to: cooperative management/performance; markets | Non-ethnographic. Interviews. |
| Rotter 2011 | n/a | Fairtrade, organic | Material related to: financial premium use | Material related to: uneven adoption of standard | Material related to: governmental legislation, incentives, and/or regulation; social infrastructure; markets | Ethnographic. |
| Roy and Thorat 2008 | n/a | GlobalGAP | Material related to: participant's targeting and self-selection; implementation costs of certification programmes | Material related to: None | Material related to: effective/non-effective adoption; production cost for certified goods | Non-ethnographic. Field surveys with questionnaire-based interviews. |
| Rueda and Lambin 2013 | n/a | Rainforest Alliance | Material related to: social premium use | Material related to: uneven adoption of standard | Material related to: participants' reception/motivation; effective/non-effective adoption | Non-ethnographic. Interviews, household survey. |
| Ryan 2011 | n/a | Fairtrade | Material related to: monitoring and auditing; financial premium use; social premium use | Material related to: uneven adoption of standard | Material related to: cooperative management/performance; participants' reception/motivation; governmental legislation, incentives, and/or regulation; markets | Ethnographic. |
| Said-Allsopp and Tallontire 2014 | n/a | Fairtrade | Material related to: training and new practices; participant's targeting and self-selection; monitoring and auditing; social premium use | Material related to: distribution of benefits & investments; elite capture | Material related to: cooperative management/performance | Non-ethnographic. Focus group discussions. |

| | | | | | | |
|----------------------------|----------------------|--------------------|---|--|--|--|
| Schelly 2011 | n/a | Fairtrade | Material related to: training and new practices; participant's targeting and self-selection; certification related services; monitoring and auditing | Material related to: uneven adoption of standard | Material related to: markets | Non-ethnographic. Semi-structured interviews. |
| Schoonhoven-Speijer 2012 | n/a | UTZ | Material related to: training and new practices; participant's targeting and self-selection; certification related services | Other | Material related to: cooperative management/performance | Non-ethnographic. |
| Schuster and Maertens 2013 | n/a | Various | None | None | Material related to: markets | Non-ethnographic. Secondary datasets and survey. |
| Sen 2009 | n/a | Fairtrade, organic | Material related to: monitoring and auditing | Material related to: elite capture | Material related to: cooperative management/performance | Ethnographic. |
| Setrini 2011 | n/a | Fairtrade, organic | Material related to: training and new practices; participant's targeting and self-selection; monitoring and auditing; implementation costs of certification programmes; financial premium use; social premium use | Material related to: distribution of benefits & investments | Material related to: cooperative management/performance | Ethnographic. |
| Shreck 2002 | Getz and Shreck 2006 | Fairtrade, organic | Material related to: training and new practices; certification related services; implementation costs of certification programmes; financial premium use; social premium use | Material related to: distribution of benefits & investments; uneven adoption of standard | Material related to: cooperative management/performance; participants' reception/motivation; markets | Ethnographic. |
| Silva-Castaneda 2012 | n/a | RSPO | None | None | Material related to: governmental legislation, incentives, and/or regulation | Non-ethnographic. Interviews, observation. |
| Smith et al 2004 | n/a | EurepGAP | None | Material related to: distribution of benefits & investments | Material related to: cooperative management/performance | Non-ethnographic. Interviews, focus groups. |
| Smith 2007 | n/a | Fairtrade, organic | Material related to: financial premium use | Material related to: uneven adoption of standard | Material related to: cooperative management/performance | Ethnographic. |

| | | | | | | |
|---------------------------------|-----|--------------------|--|--|--|---|
| Smith 2010 (Ghana) | n/a | Fairtrade | Material related to: training and new practices; implementation costs of certification programmes; financial premium use; social premium use | Material related to: distribution of benefits & investments; uneven adoption of standard | Material related to: markets | Non-ethnographic. Interviews. |
| Smith 2010 (Ecuador) | n/a | Fairtrade | Material related to: training and new practices; implementation costs of certification programmes; financial premium use; social premium use | Material related to: distribution of benefits & investments; uneven adoption of standard | Material related to: markets | Non-ethnographic. Interviews. |
| Smith 2010 (Various) | n/a | Fairtrade | Material related to: training and new practices; implementation costs of certification programmes; financial premium use; social premium use | Material related to: distribution of benefits & investments; uneven adoption of standard | Material related to: markets | Non-ethnographic. Interviews. |
| Smith 2010 (Windward Islands) | n/a | Fairtrade | Material related to: training and new practices; implementation costs of certification programmes; financial premium use; social premium use | Material related to: distribution of benefits & investments; uneven adoption of standard | Material related to: markets | Non-ethnographic. Interviews. |
| Smith 2010 (Dominican Republic) | n/a | Fairtrade | Material related to: training and new practices; implementation costs of certification programmes; financial premium use; social premium use | Material related to: distribution of benefits & investments; uneven adoption of standard | Material related to: markets | Non-ethnographic. Interviews. |
| Staib 2012 | n/a | Fairtrade, organic | Material related to: certification related services; monitoring and auditing; implementation costs of certification programmes; social premium use | Material related to: distribution of benefits & investments; elite capture | Material related to: cooperative management/performance; participants' reception/motivation; markets | Ethnographic. |
| Staricco and Ponte 2015 | n/a | Fairtrade | Material related to: certification related services | Material related to: uneven adoption of standard | Material related to: markets | Non-ethnographic. Semi-structured interviews. |

| | | | | | | |
|----------------------------|-----|--------------------------------|---|---|---|--|
| Stathers and Gathuthi 2013 | n/a | Fairtrade, Rainforest Alliance | Material related to: training and new practices; certification related services; implementation costs of certification programmes; social premium use | Material related to: distribution of benefits & investments | Material related to: cooperative management/performance; participants' reception/motivation; markets | Non-ethnographic. Formal questionnaires and qualitative checklists, interviews, focus group discussions. |
| Stenn 2015 | n/a | Fairtrade | Material related to: training and new practices; financial premium use; social premium use | Material related to: distribution of benefits & investments | Material related to: participants' reception/motivation | Ethnographic. |
| Subervie and Vagneron 2013 | n/a | GlobalGAP | Material related to: training and new practices; participant's targeting and self-selection | Material related to: distribution of benefits & investments | None | Non-ethnographic. Surveys, questionnaires. |
| Sutton 2014 | n/a | Fairtrade | Material related to: training and new practices; implementation costs of certification programmes; financial premium use; social premium use | Material related to: uneven adoption of standard | Material related to: cooperative management/performance; markets | Non-ethnographic. Interviews. |
| Terstappen 2010 | n/a | Fairtrade, organic | Material related to: training and new practices; implementation costs of certification programmes; financial premium use; social premium use | Material related to: uneven adoption of standard | Material related to: cooperative management/performance; markets | Ethnographic. |
| Trauger 2014 | n/a | Fairtrade, organic | Material related to: training and new practices; monitoring and auditing; implementation costs of certification programmes; social premium use | Material related to: distribution of benefits & investments | Material related to: cooperative management/performance; markets | Ethnographic. |
| TWIN 2012 (Côte d'Ivoire) | n/a | Fairtrade | Material related to: training and new practices; certification related services; financial premium use | None | Material related to: cooperative management/performance; governmental legislation, incentives, and/or regulation; markets | Non-ethnographic. Focus group discussions and SSIs. |
| TWIN 2012 (Kenya) | n/a | Fairtrade | Material related to: training and new practices | None | Material related to: cooperative management/performance | Non-ethnographic. Focus group discussions and SSIs. |
| TWIN 2012 (Malawi) | n/a | Fairtrade | Material related to: training and new practices | None | Material related to: cooperative management/performance | Non-ethnographic. Focus group discussions and SSIs. |

| | | | | | | |
|-------------------------|-------------------------|---------------------|--|--|---|--|
| TWIN 2013 | n/a | Fairtrade | None | Material related to: distribution of benefits & investments; uneven adoption of standard | Material related to: social infrastructure | Non-ethnographic. Interviews, group discussions. |
| Valkila and Nygren 2008 | Valkila 2009 | Fairtrade, organic | Material related to: training and new practices; certification related services; implementation costs of certification programmes; financial premium use; social premium use | Material related to: distribution of benefits & investments | Material related to: cooperative management/performance; markets | Non-ethnographic. Semi-structured interviews. |
| Valkila 2009 | Valkila and Nygren 2008 | Fairtrade, organic | Material related to: certification related services; implementation costs of certification programmes; financial premium use; social premium use | Material related to: distribution of benefits & investments | Material related to: markets | Non-ethnographic. Semi-structured interviews. |
| Vásquez-León 2010 | n/a | Fairtrade, organic | Material related to: financial premium use; social premium use | None | Material related to: markets | Ethnographic. |
| Waarts et al 2012 | n/a | Rainforest Alliance | Material related to: training and new practices | None | Material related to: participants' reception/motivation; production cost for certified goods | Non-ethnographic. Quantitative survey, focus group discussions. |
| Waarts et al 2014 | n/a | UTZ | Material related to: training and new practices; participant's targeting and self-selection | Material related to: distribution of benefits & investments | Material related to: cooperative management/performance; participants' reception/motivation | Non-ethnographic. Interviews, focus groups, quantitative survey. |
| Waarts et al 2016 | n/a | UTZ | Material related to: training and new practices; certification related services | Material related to: uneven adoption of standard | None | Non-ethnographic. Interviews, survey, focus group discussions. |
| Walsh 2004 | n/a | Fairtrade, organic | Material related to: certification related services; financial premium use; social premium use | Material related to: distribution of benefits & investments | Material related to: cooperative management/performance; markets; production cost for certified goods | Non-ethnographic. Semi-structured interviews. |
| Wilson 2010 | n/a | Fairtrade | Material related to: certification related services; financial premium use | Material related to: elite capture | Material related to: cooperative management/performance; participants' reception/motivation; markets | Ethnographic. |