

What Are We Getting from Voluntary Sustainability Standards for Coffee?

Kimberly Ann Elliott

Abstract

Demand for and supply of “sustainable” coffee (and other commodities) have grown markedly for two decades, as has the literature analyzing the effects of voluntary sustainability standards for coffee. The evidence for assessing the impacts for smallholder producers and the environment remains relatively weak, however. A relatively small number of studies use methods that allow researchers to attribute observed outcomes to sustainability certifications. This paper reviews research from the past decade on the effects of coffee sustainability schemes to see what we have learned about the impact of such schemes, and whether positive livelihood effects are mainly the result of relatively better off households choosing to participate. Overall, the available research suggests that certification schemes can be beneficial, but context matters, and the poorest, most vulnerable smallholder producers are able to comply with sustainability standards only with substantial external help.

The author is grateful to an anonymous reviewer and, in particular, to Sebastian Schienle and Ian Mitchell for their detailed and insightful comments on the paper. She is also grateful to Andrea Estrella Chong for sharing her team’s research so it could be included in the analysis. The author would also like to thank Benckiser Stiftung Zukunft for financial support of the research. This paper represents the views of the author alone and she is responsible for any errors in fact or judgment.

Kimberly Ann Elliott. 2018. “What Are We Getting from Voluntary Sustainability Standards for Coffee?” CGD Policy Paper. Washington, DC: Center for Global Development. <https://www.cgdev.org/publication/what-are-we-getting-voluntary-sustainability-standards-coffee>

The data used in this paper is available here: <https://www.cgdev.org/sites/default/files/elliott-coffee-sustainability-data.xlsx>

Center for Global Development
2055 L Street NW
Fifth Floor
Washington DC 20036
202-416-4000
www.cgdev.org

This work is made available under
the terms of the Creative Commons
Attribution-NonCommercial 4.0
license.

Contents

Comparing Standards and Verification Approaches	2
Trends in and Status of VSS Adoption in the Coffee Sector.....	4
The Evidence Base.....	5
The Impacts Identified in the Literature Reviewed.....	6
Price	6
Quality and Productivity.....	7
Income/Livelihoods	7
Working Conditions.....	9
Environmental Impacts.....	10
Assessments from Other Literature Reviews.....	10
Other Impacts of Note.....	11
Challenges and Confounding Factors	12
Who Gets Certified?	13
Summary and Conclusions.....	15
Bibliography	18
Tables and Figures.....	27

The volume of coffee produced under voluntary codes to reduce poverty and improve environmental sustainability in coffee production has been growing rapidly for two decades. The number of such voluntary sustainability standards (VSS), including private initiatives by coffee roasters and retailers such as Starbucks and Nestle, is also growing. Indeed, some coffee companies are abandoning independent initiatives that involve multiple stakeholders in favor of their own approaches, which they argue can be as effective at lower cost. Though the emphases and approaches of these initiatives vary, they generally aim to improve the economic, environmental, and social sustainability of coffee production. Yet the evidence that these initiatives deliver what they promise remains contested.

This paper follows on an earlier CGD paper documenting trends in the market for fair trade certified products and assessing issues that would influence growth in this market (Elliott 2012).¹ That paper analyzed the sources of potential costs and benefits for producers, but it did not attempt a full empirical assessment. Since then, there have been many more studies on the impacts of a variety of sustainability schemes, as well as a systematic review of evidence on the socioeconomic impact of certification schemes for agricultural commodities by the Campbell Collaboration (Oya et al. 2017).

This paper differs from those in several ways. It focuses on coffee production, and on the four major sustainability initiatives in this sector: 4C (now the Global Coffee Platform), Fairtrade (FT), Rainforest Alliance (RA), and UTZ Certified.² Coffee is an important cash crop for smallholder producers and as one of the first and most widely certified commodities is also the most widely studied (Oya et al. 2017, p.71). It is also useful to focus on a single commodity as the impacts may be quite different for other products concentrated in different geographic regions or dominated by plantation production systems. This paper also reviews evidence on the environmental, as well as economic and poverty, effects of these certification and verification schemes. And it reflects important new research efforts that were unavailable to the Campbell Review researchers.

I reviewed more than seven dozen research studies, previous literature reviews, and impact assessments commissioned to examine the implementation of these four initiatives. Many of the producers operating under these standards, especially Fairtrade, are also certified organic, and the paper distinguishes the effects of organic certification from the effects associated with the broader standards when possible. But organic standards alone are not part of the review because they do not address socioeconomic outcomes.

The paper begins with a brief summary of key elements of, and differences between, these four sustainability initiatives. It then reviews recent trends in and the current status of coffee

¹ I use the term “fair trade” for any form of fair trade labeling, while “Fairtrade” is reserved for products certified by or activities of the Fairtrade Labeling Organization International (FLO) (also Fairtrade International for short).

² In January 2018, Rainforest Alliance and UTZ Certified merged. Since the assessments reviewed here all precede the merger, the initiatives are kept separate in this paper.

certification.³ After discussing the strengths and weaknesses of the evidence base for assessing the impacts of certification, the balance of the paper summarizes the key lessons derived from the literature review. This includes impacts on the livelihoods of producers, wage workers, and the environment, as well as an assessment of the obstacles to certification and whether the most vulnerable producers are able to access and benefit from these initiatives.

Comparing Standards and Verification Approaches

The four initiatives all use compliance criteria related to economic sustainability and “good agricultural practices”; social sustainability, including decent working conditions and no child labor; and environmental sustainability. 4C and UTZ Certified focus relatively more on improving producers’ agroecological practices and helping them become more productive as the primary means to improving livelihoods and achieving increased sustainability. The Rainforest Alliance focuses relatively more intensely on decreasing the environmental impact of coffee production. Fairtrade, as implemented by the international Fairtrade Labeling Organization (FLO), is relatively more focused on changing relationships in the coffee supply chain and delivering a larger share of the benefits to smallholder producers.

Only the FLO focuses exclusively on smallholders in the coffee sector, sets a minimum price, and mandates payment of a “social premium,” which should be allocated by agreement among democratically organized producer groups.⁴ The others certify producers of any size, including larger estates. UTZ Certified requires the payment of a premium and collects data that is supposed to be available to producers, but the amount is left to negotiations between producers and buyers. Rainforest Alliance expects that producers will receive a premium but does not require it.⁵

The State of Sustainability Initiatives (SSI) Review for 2014 calculates the strength of the various initiatives based on indicators of coverage and enforceability for the economic, social, and environmental dimensions (table 1 summarizes) (Potts et al., Appendices V-VII). On all dimensions, 4C, which promotes itself as a base standard to verify (not certify) compliance with minimum standards, unsurprisingly does worst on this ranking. Rainforest Alliance does best on the social and environmental dimensions, but third on the economic dimension, mostly because it does not require a premium. FLO does best on the economic dimension, and second best on the other two. UTZ does second best on the economic dimension and third on the other two. In part, then, which code is “better” depends on one’s own priorities.

³ The Global Coffee Platform (4C) does not “certify” coffee as compliant with its standard nor does it provide a consumer-facing label. Rather it “verifies” that producers are in compliance. To simplify the presentation, however, I will generally use the term “certification” in reference to all the initiatives.

⁴ For certain other commodities where plantation agriculture is more common, such as tea and bananas, Fairtrade provides certification options for larger operation using hired labor.

⁵ The Specialty Coffee Association of America has a matrix comparing key elements of these initiatives here, <https://www.scaa.org/PDF/SustainableCoffeeCertificationsComparisonMatrix2010.pdf>, accessed July 24, 2018.

Dietz et al. (2018) create a Voluntary Coffee Standard Index that assesses each initiative, plus Fair Trade USA (and several narrower standards), on their enforcement mechanisms, as well as the breadth and strength of the economic, environmental, and social criteria included (table 1). This analysis also breaks out the separate standards that UTZ and Rainforest Alliance (as well as Fair Trade USA) apply to smallholders and larger operations. That and the inclusion of Fair Trade USA mean that the results are not directly comparable to those in the SSI report, but they are similar on the separate dimensions.⁶ Dietz et al. (ibid.) also create a composite score that covers all three dimensions plus enforcement. On that ranking, the more rigorous standards that UTZ, Fair Trade USA, and RA apply to larger operations (but not smallholders) land them in the top three spots (in that order). FLO comes in behind the other initiatives, except for Fair Trade USA's more flexible approach for smallholders.

It is interesting to note, however, that when Dietz, Grabs, and Estrella (forthcoming) compare the performance of the four initiatives on 54 economic, social, and environmental indicators for which data from Honduras is available, RA was rated ahead of UTZ. RA was "strong" on all three dimensions, while UTZ was fourth and 4C again comes out as having the weakest performance. Fairtrade, organic and conventional, occupied the second and third spots (Fair Trade USA was not included). Not surprisingly, the combination of Fairtrade and organic certification was strongest on the environmental indicators, but it was weak on the economic dimension (primarily because of negative productivity effects from the conversion to organic practices). Conventional Fairtrade producers were strongest on the social dimension and "medium" on the other two (table 6, p. 41).

Returning to the Dietz et al. VOCSI index, the enforcement dimension has fewer elements and less variation than the substantive dimensions. On enforcement, all the UTZ and RA scores (for large and small producers), as well as Fair Trade USA are between 60 and 63 (with a potential score of between 1 and 100). FLO gets the second lowest score, but it is above 50. 4C is well back with a score of 33.5 and is the only one of 14 standards assessed that falls below 50 on the enforcement dimension.

Table 2 summarizes key elements of the audit mechanisms for each of the initiatives. Given the inherent limitations of periodic, external audits, all the initiatives rely on internal management systems and detailed record-keeping as key elements of compliance. The initial certification audit is scheduled in advance so that management/group leadership can get all the documentation ready for examination. On-site visits to sampled producers or farm plots (on estates) are generally scheduled with less notice. Only FT surprise audits appear to be truly unannounced. 4C and UTZ allow no more than 72 hours of notice of "unannounced" audits, which can occur in response to complaints or receipt of information suggesting a problem could exist. 4C says only that unannounced visits can occur "when applicable," and does not provide detail on what that means in practice.

⁶ Fair Trade USA broke away from FLO in 2011 over the issue of certifying coffee estates and individual producers not organized in coops. The other elements of the US standard are generally based on the FLO.

All the initiatives try to balance the credibility of the process with the costs of certification. Only UTZ requires an audit every year, while RA and FT use three-year cycles, with less extensive surveillance audits in between. For small producer groups, FT uses a six-year cycle. 4C requires only one audit every three years and self-assessments in the interim. Unannounced audits can be used when noncompliance is suspected.

All the initiatives base their sampling methods for site visits on some version of the “square root” rule, which bases the sample size on the square root of the number of all group members. The weakness of that rule is that the share of the group subject to physical audit declines as the total number of members increases. 4C, implemented by Coffee Assurance Services, is the weakest, requiring that only *half* of the square root number must be audited. FLO-Cert sets minimum numbers of workers to be interviewed and producers to be visited for groups of different sizes. Those numbers tend to be a bit more than the square root of the upper end of each range up to groups of 1,000 members or more. For groups larger than 1,600, the number sampled could fall to below the square root threshold under the FLO guidance.

All the initiatives are members of the ISEAL Alliance, which provides guidance on sound verification and labeling practices. ISEAL does not set a minimum for sample sizes for audits and has noted that the square root rule is convenient, but has no scientific basis and is problematic for large groups.⁷ They recommend using risk-based assessments to set the number and to select the sample, though they also note that the selection should be random to ensure that it is representative. All the initiatives embrace risk assessments for audits to some degree. FT will increase the frequency of surveillance audits based on risk, while RA calls for increasing the audit sample by up to 1.4 times the square root number if necessary to address risk levels. It is not clear how any of the initiatives implement these safeguards in practice.

Trends in and Status of VSS Adoption in the Coffee Sector

Certification (or verification) by the four main coffee initiatives, as measured by the number of hectares and the volume of production covered, has grown rapidly over the past two decades. Potts et al. (2014, p. 161) report that the share of coffee produced globally under at least one of seven sustainability standards (including Nespresso AAA, Starbucks C.A.F.E., and organic) grew from 15 percent in 2008 to 40 percent in 2012. The authors adjust the figures to avoid double-counting due to the fact that many coffee estates or producer groups hold more than one certification. Figure 1 shows the trends in total coffee production that is certified by each initiative. Though the rapid growth in 4C coffee dominates, certified coffee under the other initiatives also doubled or tripled over this period.

Table 3 provides a snapshot of each of the initiatives in 2015 (or the most recent year for which data is available). The notable thing here is the relatively small share of certified production that producers are actually able to sell on certified terms. Over the period shown

⁷ See the blog post by [Patrick Mallet](#), ISEAL Director of Innovation, accessed November 2017.

in figure 1, Fairtrade certified sales have been relatively steady at around 30 percent of certified production while UTZ increased the share from around a quarter to a third more recently. 4C, however, showed the strongest gains, increasing the share of certified sales from around 5 percent to around 20 percent in recent years. Overall, the share of certified production sold as such for these four initiatives was 25 percent in 2014; excluding 4C it was a third.

The other notable feature of the table is the high profile of Latin America in certified coffee production. Overall, including private and organic initiatives, Brazil accounted for 43 percent of “standard-compliant” coffee production in 2012, while Colombia was second with 18 percent, and Peru, Mexico, and Central American added another 13 percent for a total of 74 percent. Vietnam was third overall, with 16 percent of certified production, most of it under 4C and the balance under UTZ (Potts et al., pp. 178-79).

The Evidence Base

Despite a substantial increase in both the quantity and quality of VSS studies, it is still difficult to draw general conclusions about the impacts of certification on poor coffee producers, their workers, or the environment. In addition to the methodological weaknesses in the majority of studies, the research also underscores the degree to which context matters, so that what works under one set of circumstances may not in another (Oya et al. 2017, pp. 183-84; Giovannucci and Potts 2008, p. ix).

A key problem is that there are still too few studies well designed enough to draw causal inferences about the impact of VSS. There are very few studies that collect data over time and fewer yet that have baseline data showing the condition of producers prior to certification. Unlike the early days of fair trade research, many researchers now do include similar, uncertified producers (a counterfactual) in their analysis to control for broader trends in markets and with weather or other factors affecting production and prices. But case studies and household samples are rarely selected randomly and control (uncertified) and treatment (certified) groups often have significant differences. So differences in observed outcomes could be due to omitted variables, rather than certification.

More researchers are using “quasi-experimental” methods that attempt to control for the bias potentially introduced if producers choosing certification are qualitatively different from those who do not.⁸ These methods include “propensity score matching,” in which researchers use statistical methods to match treatment group members with the closest matches from the control group. But that method is limited to observable differences and there could be unobservable differences—certified producers are more entrepreneurial, for example—that confound the results. Another method is to replace certification in regression analyses with an “instrumental variable” that is correlated with certification, but not with the outcome measures of interest. This method is less common, probably because it is so

⁸ COSA (2013) and Neilson and Toth (2016) report that they are working on randomized control experiments of the effects of certification, which is the preferred methodology.

difficult to identify good instruments. Overall, even for the 15 coffee studies that the Campbell Review authors deemed rigorous enough for inclusion in their statistical analysis, 14 were assessed to have a medium or high risk of bias. For this review, I added seven additional cases from two studies that use these quasi-experimental methods (Akoyi and Maertens 2016, Dietz et al. forthcoming).

To compile the list of studies examined for this review, I started with the Campbell Review (Oya et al. 2017) and followed their assessment of the relative quality of the studies. The Excel workbook that accompanies this paper has five separate worksheets for different categories of study types. The first includes 22 cases from 15 studies that met the Campbell Review's rigorous standard for inclusion in its quantitative analysis of the impacts of standards; the second summarizes impacts reported for 18 cases from 17 studies included in the Review's qualitative analysis (note that parts of Riisgaard et al. 2009 are also included in the quantitative review); the third has 25 cases from 17 other studies that included at least a weak counterfactual; the fourth summarizes three qualitative studies with weak or no controls; and 16 cases from nine impact assessments or reports commissioned by the initiatives (and not included in the Campbell Review) are presented in the last worksheet.

Of the cases summarized in the workbook, 49 examine Fairtrade (roughly half and half conventional and double-certified), 32 (including 11 commissioned case studies) address UTZ, 13 look at RA, and only four, three of them commissioned by the initiative, examine 4C. Latin America, not surprisingly, dominates, with 59 cases, followed by Africa with 34, and Asia with 11 (mostly Vietnam).

The Impacts Identified in the Literature Reviewed

Many analyses of voluntary coffee standards assess the impact on prices and net income or other measures of producer livelihoods; some also look for effects on product quality, yield/productivity, working conditions and child labor, or the environment. What follows are observations based on the weight of the evidence from the studies reviewed, qualitative as well as quantitative. Given the variation in conditions across the case studies, as well as differences in methodological rigor, readers should not interpret these observations as generalizable.

Price

Most studies report on price effects and most find that they are positive. Most of the available data comes from studies of Fairtrade, which sets a minimum price for certified coffee, and they show that the price premium producers receive varies inversely with market price. This is not surprising since, under Fairtrade terms, buyers will only be paying the (modest) social premium when the market price is above the minimum price. There is less available data on the size of any premium associated with the other initiatives and it is not possible to know whether they vary in the same way, though it seems likely that they do. Any premium on the overall average price for coffee that any individual producer receives is correlated positively with the share sold as certified. Becchetti et al. (2015) confirm this in

their recent review of the quantitative evidence on Fairtrade impacts across commodities, including coffee. They conclude that the main benefit comes from the insurance effect of having a price floor. Wilson (2010) similarly finds that the FT price floor provided a crucial safety net during the price crisis of the early 2000s but he nevertheless concludes that certification cannot overcome the marginalization of very poor smallholders (see also Bacon et al. 2008).

Some studies of RA and UTZ also report higher prices, but the premia are usually attributed to improved product quality and are small. Studies of all the initiatives find that the price premium at the farm gate is relatively small even for coffee sold on certified terms because the coop or other management entity deducts costs related to certification and the administrative burden of compliance. Organic and FT double certified coffee typically receive higher premia (but also entail higher costs as discussed below). Only two studies report negative price effects and both have methodological problems. Only one of four studies find a (small) positive price effect for 4C.

Quality and Productivity

Relatively few studies (especially independent ones) report on product quality and those that do almost always report that quality improves after certification or is higher among certified producers. But that is often due to training, improved processing facilities, or buyer requirements that are not directly related to certification and for which the researchers do not adequately control.

The results on productivity are, unsurprisingly, mixed for all the initiatives. The results are somewhat more positive for UTZ, which is not surprising given the emphasis on good agricultural practices, and somewhat more likely to be zero for Fairtrade, which emphasizes the trading relationship and prices over agricultural practices. Studies of the Rainforest Alliance rarely report productivity effects, but those that do tend to find positive effects.

Studies that assess organic certification (usually double-certified as Fairtrade) tend to find that it has negative yield effects, though the context very much matters. Some studies (for example, Jaffee 2007) find that relatively poorer, more marginalized producers who were passively organic (not using inputs because they could not afford them) initially could gain from adopting organic practices. But that was not the case in Ethiopia (Minten et al. 2015). Other studies report that yields were negatively affected, at least initially, when producers gave up inorganic chemicals after adopting organic practices (Barham and Weber 2012; Valkila 2009, Hoebink et al. 2014, Abarca-Orozco 2015, Nelson et al. 2016).

Income/Livelihoods

It is difficult to summarize the findings on household welfare because studies report on a wide variety of indicators and often do not provide information on the costs of certification, meaning they do not provide information on the *net* income effects. Certification costs are important because they are a frequent reason that uncertified producers say they choose not

to participate. Many studies also report only the impact on income from certified production, which may be different from the impact on overall household income if producers shift additional resources into coffee and away from other income-producing activities.

Several credible studies find positive impacts on coffee income (though a mix of gross and net). Others find positive impacts on other indicators of household welfare, such as assets (savings, credit, livestock, or other farm investments), food security, or education. Very few find credible evidence of overall improvement in net household incomes.

While only a few studies find negative effects (one for 4C, three for FT, three for FT/organic double-certified), many find no or mixed effects, especially when researchers consider costs and assess net effects. Several studies also cite the increasing pressure on FT coops to be double-certified and to adopt organic practices as increasing costs (and reducing yields) by enough that net effects are often nil or sometimes even negative (Barham and Weber 2012, Valkila 2009). While organic production lowers some input costs, it raises labor costs and the lower yields can contribute to weak or negative net benefits when premia are low.

Another obstacle to larger net welfare benefits is the low share of certified production that is actually sold as such. De Janvry et al. (2014) have detailed data for FT-certified coops in Central America demonstrating how free entry into certified markets reduces the price benefit. Yet these coops must still cover the costs of coop management and certification, which further reduces or even eliminates the amount of price premium that trickles down to producer.

Analyses of 4C find few benefits, but also low costs of certification and compliance. Studies of UTZ, RA sometimes report positive income effects as a result of producers reducing their costs, either through overall better farm management, or by reducing inputs. Studies focused on RA are somewhat more likely to find positive effects, while studies of FT and UTZ either have mixed results or are roughly split between positive or no effect on broader welfare measures.

Since certification is not costless, some studies also try to compare the income effects of voluntary standards compliance with alternative investments that may be available to households. Weber (2011) conducts a detailed analysis of sources of household income in Mexico and finds that, while FT/organic certification raised (gross) incomes, the income effects of investing in education or migration, or by government subsidy programs for the poor (Oportunidades) and for farmers (Procampo) swamped the effects of certification (see also Barham et al. 2011).

Dietz et al. (forthcoming), who systematically compare all four initiatives in Honduras, find no effect on gross coffee incomes of FT certification or 4C membership, a negative effect for FT/organic double certification, and positive effects for adherence to RA and UTZ. Overall, the studies reviewed here report mixed effects on producer livelihoods and that the effects of certification depend on context (Oya et al. 2017, pp. 183-84). Among the issues

that multiple studies identify as affecting outcomes is the effectiveness of coops in representing, managing, and communicating with producers (Jena et al. 2012; Abarca-Orozco 2015; Babin 2012; Fraser et al. 2014; Milford 2014; Staib 2012). Many researchers also find that having a strong relationship with the buyer is associated with more positive outcomes (Abarca-Orozco 2015; Hoebink et al. 2014). Studies focusing on more marginalized communities argue that where yields and production levels are low, the price premia associated with certification are not enough to overcome those fundamental problems and certification has little effect on poverty (Mendez et al. 2010, Saenz Segura and Zuniga-Arias 2008; Barham et al. 2011, Beuchelt et al. 2009, Beuchelt and Zeller 2011).

Working Conditions

Some studies consider effects on child labor, health and safety issues for producers, wages, or living conditions for hired labor. The International Labor Organization's core standards of freedom of association, forced labor, and nondiscrimination rarely attract much if any attention. The SOAS report for DfID is the only study to focus specifically on the effects of FT for hired labor, noting that there is more such labor than the FT focus on smallholders recognizes. This study finds little or no benefits for hired workers in Africa, especially in Ethiopia. But closer examination suggests that much of that result could be explained by the inclusion of large estates in the counterfactual (noncertified) samples. When smallholders are compared to smallholders, there are some modest positive effects of FT certification even in this study.

In a study using detailed survey data from Honduras, Dietz et al. (forthcoming) systematically analyze indicators for the social dimension (as well as economic and environmental impacts). They find that RA and FT are strong (by their definition) on this element, while FT/organic is ranked in the middle, and UTZ and 4C are weak. Beyond this study, results from others on the labor and social dimension are mixed and quite often the finding is that certification has little or no effect on labor standards compliance or working conditions. While, Dietz et al. (ibid.) find that FT performs well on social criteria, most of the FT-focused studies (except SOAS) emphasize the livelihood effects on small-holders and pay little attention to working conditions. In contrast to Dietz et al., the impact reports commissioned by UTZ find improved working conditions in several cases, but these studies are weak methodologically.

A common finding in studies that report on social impacts is that workers receive more training on health and safety issues, such as use of pesticides and other chemicals. But most studies also find that the protective gear called for in the training is often unavailable. Rainforest Alliance impact assessments report improvement, but also continued problems with noncompliance in this area; some RA-certified producers report being better able to attract returning seasonal labor as a result of improved conditions (Hughell and Newsom, Milder and Newsom, and Tuinstra and Deugd).

An independent impact assessment commissioned by UTZ (Garcia et al.) found less improvement than expected (e.g., on safety training and use of written contracts) and that where there were improvements, they were not sustained. Other UTZ assessments (COSA

2013; Kuit et al. 2013) found no change in conditions in Vietnam because performance on social criteria was high before certification.

Environmental Impacts

Most studies focus on the economic impacts and only a few of the most rigorous studies address environmental impacts. Among the studies that do report on environmental outcomes, the most common positive effects are reduced or improved agrochemical handling and improved water and/or waste management and protection. Some Rainforest Alliance studies also find improved biodiversity or tree cover (Milder and Newsom, Rueda and Lambin, Riisgaard et al.). Bacon et al. (2015, pp. 392-93) argue that the increase in the FT premium for double-certified (FT and organic) in 2011 helped more producers to become organic and thereby contributed to improved environmental outcomes. Dietz et al. (forthcoming) find the performance of RA and FT/organic in Honduras to be strong, while that of FT and UTZ are in the middle, and 4C is, again, weak. Overall, more studies of RA, which focuses on environmental sustainability, report on environmental effects and find some benefits.

Most studies, however, report on inputs (increased training) or outputs (better water waste management) not on actual outcomes. Hughell and Newsom (2013) is one exception, done for Rainforest Alliance, which looks at scientific measures of water and soil quality. There were signs of improved water quality in one study area but not the other, and they found no differences in soil quality. That, unfortunately, suggests that even relatively strong standards produce limited benefits in practice.

Overall, producers seem to adopt and maintain practices that either improve profitability, such as reducing input costs through more efficient fertilizer use, or have clear health benefits, such as water and waste management.

Assessments from Other Literature Reviews

As a supplement to this qualitative review, table 4 summarizes what two other literature reviews concluded about the general direction of potential impacts from the application of sustainability standards (ITC 2011, Kuit and Waarts 2014). Table 5 summarizes the quantitative analysis of effects from the Campbell Review. There is, of course, substantial overlap in the underlying studies reviewed.

Table 4 generally confirms the conclusion that the results of certification are mixed, but that few studies find negative effects. And, as above, the most common positive impact reported is on prices, while the net effect on incomes is more often neutral or mixed. Interestingly, the ITC review finds more positive effects for producer livelihoods, which appears to reflect some of the studies discussed above that find positive effects on assets, education, or food security, if not total income. As both these studies focused on the impacts on producers, impacts on workers and the environment are underrepresented.

The Campbell Review (Oya et al. 2017) selected 43 studies that met their methodological and reporting requirements for using statistical methods to estimate overall, average economic impacts from the adoption of sustainability standards. But not all the studies reported on all the effects they analyzed. So for any given effect, the number of studies is generally quite small and there is a high degree of heterogeneity underlying the average effect. They also reported results for specific commodities and initiatives where possible. Table 5 reports the results and the extent of the heterogeneity in the data is reflected in the final column of the table, which shows specific results for studies of the coffee sector. (Note that the effect size measure for the coffee studies is a standardized, technical measure that allows the authors to compare the results across the studies. It is not a percentage effect, which the authors do not report for specific commodities.)

Overall, this analysis suggests some modest benefits from certification (all commodities, initiatives), but the results are not robust and the authors emphasize the “wide range of contextual factors” that influence the impacts from certification initiatives. The Campbell Review’s quantitative analysis thus primarily serves to underscore the weakness of the evidence base for rigorously assessing the effects of VSS in coffee. The authors conclude (pp. 113-14):

In summary, we find that the available evidence does not give a clear picture of the impact—or lack thereof—of certification schemes. . . . While the evidence in hand points largely towards findings that are not statistically significant, the evidence base is also too thin in most cases to have great confidence in these findings.

Examining an earlier and even thinner evidence base (for all initiatives and commodities), Blackman and Rivera (2010) concluded “that empirical evidence that sustainable certification has significant benefits is limited.”

Other Impacts of Note

A number of studies cite access to training and capacity-building, markets, and occasionally credit or other financing as benefits that go beyond the direct economic effects. Many studies also point to the role that good coop management plays in realizing positive effects from certification (Jena et al. 2012, Babin 2012, Beuchelt et al. 2009, Francesconi and Ruben 2014, Schoonhoven et al. 2014). Other studies mention the important role that steady, supportive buyers can play (Abarca-Orozco 2015 for FT, Hoebink et al 2014 for UTZ). Donovan and Poole (2014b), however, note that even with a committed buyer and a well-managed coop, there were problems in getting information about and adoption of improved practices down to the household level.

On the role of training relative to certification in promoting more sustainable production, the DE Foundation and the Dutch Ministry of Economic Affairs commissioned Kuit Consultancy (2013) to assess UTZ certification in Vietnam. The authors found that certification increased access to training but that what UTZ normally provides in the context of certification had little impact on productivity or household incomes. The study concluded that the amount of training needed to have an impact was of an intensity that would likely

require ongoing donor support to maintain. They also noted, however, that Vietnamese producers are generally already among the more productive and that UTZ certification and associated training might have more impact in less productive areas. In another impact assessment of 4C implementation in Uganda (Kuit et al. 2016), however, the authors did not find that to be the case.

Despite the fact that a central aim of FT is to change the relationship between producers and buyers, and to promote more stable relationships, it is striking that only a few studies explicitly address this. Reynolds (2009) collected information from coops in Mexico and Peru and concluded that mainstream, corporate buyers, such as Starbucks and Nestle, were failing to create these relationships, “unlike other Fair Trade buyers” (cited in Dragusanu et al. 2014, p. 228). But there is little systematic information available about these supply chain relationships.

Relatively few studies address gender issues. Chiputwa and Qaim (2014) conclude that improved nutrition in the households they studied in Uganda could be due to the increased role of women in decisionmaking, which might have been due to an increased need for their labor with certification (FT, organic, and UTZ, without distinction among them). Smith (2015) finds that FT certification has done relatively little to impact gender issues because so few women in affected countries own land. She notes that the social premium could be invested in ways that have disproportionate benefits for women, but there is little information on how producer organizations use these funds (p. 416).

On 4C, Kuit et al. (2016) find that producers receive more training and appreciate it, but few changes in practices or results are detected. As noted for the Vietnam case, this could be because productivity is already high. But the result of no improvement, despite low productivity, is more puzzling for Uganda. Nevertheless, most producers plan to stay with the initiative because costs are low. This is similar to the conclusion of Dietz et al. (forthcoming) that 4C is the “weakest” standard but also the most accessible.

Challenges and Confounding Factors

A commonly mentioned problem is that producers in groups or coops have low awareness of certification or what it means. That makes it difficult for them to know whether or what benefits they may be receiving. With Fairtrade, there is often little understanding of the social premium and how it is allocated, despite the nominal requirement that the decision be made by the coop as a group. The lack of transparency can in some cases foster distrust and undermine coop effectiveness. Overall, many studies conclude that the administrative competence, transparency, and representativeness of coop management are closely related to the degree to which certification benefits smallholders (see above).

Many smallholder coops, and sometimes other coffee producers, receive training and technical or financial assistance from national coffee associations, NGOs, donors, or other outside actors to help them comply with standards initially and, sometimes to cover direct certification costs. That makes it difficult to disentangle the effects of certification (K.P.M.G. 2013). Garcia et al. (2014) discuss the role of the Colombian coffee growers’ association,

which has been particularly active in recent years, in their assessment of UTZ certification. Rueda and Lambin (2013) come to similar conclusions regarding the role of the association in RA certification in Colombia.

Hoebink et al. (2014) point to the role of having a steady buyer in providing incentives for producers to maintain improved agronomic practices, which is not necessarily related to certification. Bitzer et al. (2012) also discuss a case of RA and UTZ in Peru where producers were more “beneficiaries” than partners with NGOs and buyers in attaining certification. In those cases, the producers often did not manage to build sustainable, effective coops. One group that organized on its own initiative was relatively more successful overall during the study period.

Fewer studies of FT mention decisions to drop certification, perhaps because the social premium is often used to cover the costs. It is somewhat more common to find studies of RA and UTZ, especially in lower income countries in Africa, reporting that groups either drop certification because costs exceed (perceived) benefits, or will have to do so if NGOs or others do not continue to pay certification costs (Riisgaard et al. 2009).

Who Gets Certified?

RA, UTZ, and 4C certify farms of all sizes, including large estates. In those cases, producers often report they were already compliant with the required criteria, or close to it, and they sought certification for recognition of that fact, or for new market opportunities when buyers preferred certified suppliers (Grabs et al. 2016). In other cases, some studies note that buyers, NGOs, or others promoting certification would initially seek out more advanced, somewhat larger producers to ensure adequate supplies and the success of the certification (Rueda and Lambin 2013, RA).

Where detail about households is reported, certified producers often appear relatively better off on one or more wealth or welfare dimensions than their uncertified neighbors. Because most of the studies reviewed do not have pre-certification baseline data, it is difficult to know whether those producers were able to get certified because of those advantages, or whether welfare increased as the result of certification. Given the costs of certification, it is probably safe to assume that even FT, with the focus on smallholders, does not attract the poorest or most marginalized producers around.

Angula (2010) examines factors influencing participation in UTZ certification in Uganda and concludes that the decision to participate or not and the intensity of that participation depend on producers having a minimum level of assets. He finds that even when certification costs are covered and a price premium is offered, some producers will not have enough land, labor, access to credit, or other assets to meet the quality requirements of coffee buyers seeking certified suppliers.

Where certified farmers are noted as being smaller, less educated, or poorer, it is often in Africa, or the poorer countries of Central America (Nicaragua, Honduras) and they often become certified only with the help of NGOs (Ruben et al. 2014; Jena et al. 2012; Beuchelt

et al. 2009; Valkila 2009). In those cases, the evidence compiled to date raises questions as to whether the benefits are large enough to sustain standards compliance over time without continued external assistance. Going further, the DE Foundation (which commissioned the assessment by Kuit et al. 2013, p. vi) concluded that certification may not be commercially viable for groups of small, low-productivity producers where any premium earned is on a volume of production that is simply too small to cover the fixed costs of certification.

Grabs et al. (2016) use spatial analysis to analyze the demand for and supply of VSS initiatives in Central America and Colombia. They conclude that overall, certification tilts toward larger farms and groups that can tap into economies of scale, and ensure quality that is high enough to satisfy the specialty coffee market. Not surprisingly, they find that RA and UTZ are more likely to certify larger individual farms and estates. They also find that RA and UTZ supply chains tend to be “buyer-led,” meaning that roasters or exporters encourage suppliers to get certification as part of the buyer’s strategy to expand markets or manage risk. They find that FT certification is more likely to be supply-driven, with producer groups seeking higher prices. FT-certified coops are also relatively more likely to hold multiple certifications.

Tuinstra and Deugd (n.d., p.1) conducted a detailed analysis of RA certification in Latin America to try and determine whether certification costs were an obstacle to expansion. Despite not getting data on the price premium as they had hoped, and assuming a relatively high one, their analysis conclusion underscores the challenges that producers face in making certification pay off:

[I]mplementation costs do not constitute an entry barrier and do yield positive returns, *provided adequate guidance is given on cost-efficient technology, farm productivity and satisfactory price premiums are paid* (emphasis added).

When the Sustainable Agriculture Network announced its decision to no longer participate in certification schemes, Andre de Freitas explained that there were limits to potential benefits from certification due to complexity, a low benefit to costs ratio, difficulties in scaling it, and overall limited effectiveness. Thus, de Freitas concluded:

The above limitations mean that certification will work for farms that are already reasonably well-managed, have access to resources, have access to markets that are able to better value their products, and encounter fairly well-functioning local governance structures. These conditions are very specific and are not the reality most farmers in the world live in.⁹

⁹ See the commentary by [Andre de Freitas](#) in “It’s Time to Recognize the Limits of Certification in Agriculture,” *Mongabay*, November 16, 2017.

Summary and Conclusions

The key conclusions with respect to the evidence base for the impacts of coffee certification include the following:

- Though growing, the body of research from which firm conclusions about whether and how VSS improves economic, social, or environmental outcomes is still thin.
- Fairtrade remains overrepresented while there are hardly any studies of 4C.
- Latin America is still the focus of most studies, as the bulk of certified coffee comes from there; Africa is attracting more attention, but Asia is underrepresented in the research.

With respect to impacts:

- Overall the available evidence suggests that certification can have modest, positive effects and researchers find relatively few negative impacts.
- There is evidence that certification improves prices for farmers and can increase income from coffee, but, for a variety of reasons, these benefits may not lead to increases in overall net household income.
- The costs of compliance and certification absorb much of any price premium and producers discern little direct benefit when the share of coffee sold as certified is low.
- Low awareness among producers that they are certified, or what that means, is an obstacle to effective implementation.
- Studies focusing on worker wages are few and find little or no effect; some studies report improved training for workers in safe use of agrochemicals, but evidence of sustained implementation of worker protection is weaker.
- A number of studies nevertheless report that certification is associated with improved handling and use of agrochemicals, and improved water use and protection.
- The increasing trend toward organic certification, along with FT certification, can reduce the use of inorganic chemicals, but also negatively affect yields and revenues (at least in the short run).

With respect to contextual factors, qualitative research suggests:

- Effective, transparent producer organizations play a critical role in capturing and spreading the benefits of certification to smallholders.
- Steady, supportive relations with buyers are helpful.
- Farmers appreciate training linked to certification, but the impact on sustained implementation of improved practices and, ultimately, on productivity or quality is difficult to detect.
- The provision of financial and technical assistance to help smallholder producer organization get certification makes it difficult to disentangle the effects of certification itself.

With respect to who benefits:

- Larger producers and estates seeking certification, or especially verification by 4C, report that they are generally already in compliance with standards and make few changes as a result.
- Certification can help smallholder producers, at least modestly, but they are unlikely to be the poorest because those producers often cannot achieve the quality increasingly demanded by buyers and do not have the land, labor, educational or other resources to comply with standards, or make certification worthwhile.

The bottom line is that certification is only *commercially* viable if producers benefit perceptibly. That means that agroecological and labor standards will only be implemented and maintained if they improve productivity, or if there is a price premium large enough to cover the costs. The studies suggest that the FT minimum price will be helpful to producers if prices again fall through the floor for a sustained period of time, and that the social premium is helpful in covering the costs of certification. Research on the Rainforest Alliance suggests that it might do more to improve environmental outcomes, but those studies are also more likely to address environmental issues than studies of other standards. Overall, the evidence on impacts is mixed and usually finds modest effects at best. The poorest, most marginalized smallholder producers have trouble getting certified without external assistance.

The available research is not sufficient to point to any of the initiatives as being clearly “better” than the others. Since they emphasize different priorities (prices/income, quality/productivity, environmental sustainability), which standard is “better” also depends on the buyer’s or consumer’s preferences. The exception is 4C, which positions itself as an entry-level standard and does not purport to match the others in terms of the obligations it imposes on producers. The Campbell Review (Oya et al., 2017, p. 10) found that outcomes on yields and income were more mixed for Fairtrade (across all commodities) than for other initiatives, but that could simply be a result of the larger number of FT studies available. The index developed by Dietz et al. (2018) ranks the various coffee initiatives according to the strength of their standards and enforcement on paper. But their analysis of the four main initiatives in practice in Honduras suggests a different ranking, highlighting the difficulties in comparing the initiatives.

To the degree there are benefits to certification, a key obstacle to increasing or extending those benefits is the fact that the supply of certified products is so much larger than the demand. Producers must invest time and other resources up front to obtain certification, but often without a guarantee that they will be able to find buyers that value the extra effort. When only a portion of certified coffee receives a premium price, the additional revenue may not be enough to cover the additional costs involved in certification. To better assess the extent of this problem, it would be useful for the initiatives to release more information about turnover among certified operations and how long they tend to maintain their certification.

Thus, while sustainable coffee certification has seen rapid growth, there are challenges from a number of angles. In recent years, a number of large brands and retailers have opted to rely

on their own, internal sustainability efforts because of concerns about high costs and limited perceived benefits from third party certification (so far mainly among those selling tea and chocolate).¹⁰ Whether that trend turns out to be a viable alternative will depend on the rigor—and transparency—with which companies implement these private efforts, and promote them to their customers. From the other direction, the decision by the Sustainable Agriculture Network, which had previously collaborated with RA, to withdraw from certification efforts, and the decision by Rainforest to merge with UTZ, suggest that at least some of these initiatives are having trouble responding to the demands for greater efficiency and effectiveness. There will remain a core group of consumers that will look for assurances that their coffee (and other commodities) is produced sustainably. How much that group will continue to grow is the big question.

¹⁰ See “[How fair is our food? Big companies take reins on sourcing schemes](#),” *Reuters*, September 3, 2017.

Bibliography

Impact studies commissioned by VSS organizations (or published on their websites)

4C Association

- 4C Association and Centre for Regional Entrepreneurial and Coffee Studies (CRECE) (2016). Impact Evaluation of Implementing 4C Entry Level Standard in Colombia, 4C summary plus management response. [Link to full study broken]
- Kuit Consultancy (2016). 4C Impact Study Phase 2: Estimating the impact of implementation of the 4C entry level standard in Uganda and Vietnam (plus management response).

Fairtrade International

- Center for Evaluation (2012). Assessing the Impact of Fairtrade on Poverty Reduction through Rural Development: Final Report, Fairtrade Impact Study. Commissioned by TransFair Germany and Max Havelaar of Switzerland.
- Fairtrade International (2017). Monitoring the Scope and Benefits of Fairtrade, 8th edition.
- Nelson, Valerie, and Barry Pound (2009). The Last Ten Years: A Comprehensive Review of the Literature on the Impact of Fairtrade. Natural Resources Institute (NRI), University of Greenwich.
- Nelson, V., J. Haggar, A. Martin, J. Donovan, E. Borasino, W. Hasyim, N. Mhando, M. Senga, J. Mgumia, E. Quintanar Guadarrama, Z. Kendar, J. Valdez, D. Morales (2016). 'Fairtrade coffee: A study to assess the impact of Fairtrade for coffee smallholders and producer organisations in Indonesia, Mexico, Peru and Tanzania. Natural Resources Institute, University of Greenwich, Chatham, UK (with management response).
- True Price (2017). Assessing Coffee Farmer Household Income. Fairtrade International.

Rainforest Alliance

- Milder, Jeffrey C., and Deanna Newsom (2015). SAN/Rainforest Alliance Impacts Report: Evaluating the Effects of the SAN/Rainforest Alliance Certification System on Farms, People, and the Environment. Rainforest Alliance and Sustainable Agriculture Network.
- Whelan, Tensie, and Deanna Newsom (2014). Sustainable Coffee Farming Improving Income and Social Conditions Protecting Water, Soil and Forests. Rainforest Alliance.
- Hughell, David, and Deanna Newsom (2013). Evaluating the Results of Our Work: Impacts of Rainforest Alliance Certification on Coffee Farms in Colombia. Rainforest Alliance and Sustainable Agriculture Network.
- Tuinstra, Alexandra, and Michelle Deugd (n.d.). Rainforest Alliance Certification in Coffee Production: An analysis of Costs and Revenues in Latin America 2010-11. Rainforest Alliance.

UTZ CERTIFIED

- UTZ CERTIFIED (2017). UTZ Response to Campbell Review of Certification Schemes.
- UTZ CERTIFIED (2016). UTZ: Combining insights from UTZ monitoring data with findings from impact studies.
- UTZ CERTIFIED (n.d.). The UTZ Theory of Change: Influencing factors and unintended results.
- BSD Consulting and Ibi Êté Consultoria Ltda. (2015). Effects of UTZ certification according to coffee farmers in Brazil. Sao Paulo (commissioned by and available, along with management response, on UTZ website).
- García, C., J. García, G. Ochoa, J.C. Mora, and J.F. Castellanos (2014). Impact Evaluation of UTZ Certified Coffee Program in Colombia (2008-2012). CRECE. Manizales, Colombia (commissioned by and available, along with management response, on UTZ website).
- COSA (2013). Vietnam coffee: A COSA Survey of UTZ Certified Farms. Philadelphia: Committee on Sustainability Assessment (available, with management response, on UTZ website).
- Centro de Investigaciones Humanismo y Empresa (2010). Social, Economic and Environmental Results of UTZ Certification: Case studies on UTZ CERTIFIED coffee farms in Asia, Africa and Latin America. Universidad del Istmo, Guatemala. UTZ CERTIFIED.

Third party studies (2008-2017)

- 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.
- Akoyi, Kevin Teopista and Miet Maertens (2016). "Private Sustainability Standards in the Ugandan Coffee Sector: Empty Promises or Catalysts for Development?." Bioeconomics Working Paper Series #2016/3. Leuven, Belgium: Department of Earth and Environmental Sciences, University of Leuven.
- Angula, Martin Ndinomupya (2010). Determinants of sustainable coffee marketing channel choice and supply response among organic and utz certified Smallholder Farmers: evidence from Uganda. Michigan State University. Agricultural, Food and Resource Economics.
- Anteneh, A., et al. (2014). "The More the Better? The Effects of Multiple Certification on Smallholder Coffee Farmer Livelihood: Evidence from Southern Ethiopia." In Hoebink, P., Ruben, R., Elbers, W., and van Rijsbergen, B., ed. The impact of coffee certification on smallholder farmers in Kenya, Uganda and Ethiopia. Report prepared by Centre for International Development Issues Nijmegen, Radboud University Nijmegen, for Solidaridad.
- Arce, A. (2009). Living in Times of Solidarity: Fair Trade and the Fractured Life Worlds of Guatemalan Coffee Farmers. *Journal of International Development*, 21, pp.1031-1041.
- Arnould, E.J., A. Plastina, and D. Ball (2009). Does Fair Trade Deliver on Its Core Value Proposition? Effects on Income, Educational Attainment, and Health in Three Countries. *Journal of Public Policy & Marketing*, 28(2), pp.186-201.

- Babin, N. (2015). The coffee crisis, fair trade, and agroecological transformation: impacts on land-use change in Costa Rica. *Agroecology and Sustainable Food Systems*, 39(1), pp.99-129. [Based on PhD]
- Bacon, Christopher M., et al. (2008). *Confronting the Coffee Crisis: Fair Trade, Sustainable Livelihoods and Ecosystems in Mexico and Central America*. Cambridge, MA: MIT Press (Nicaragua chapter).
- Bacon, Christopher M., Robert A. Rice, and Hannah Maryanski (2015). "Fair trade coffee and environmental sustainability in Latin America." In Reynolds, L.T., and E.A. Bennett, editors. *Handbook of Research on Fair Trade*. Edward Elgar Publishing.
- Barham, B.L., and J.G. Weber (2012). The Economic Sustainability of Certified Coffee: Recent Evidence from Mexico and Peru. *World Development*, 40(6), pp.1269-1279. [assessed as critical risk of bias]
- Barham, B.L., M. Callenes, S. Gitter, J. Lewis, and J. Weber (2011). Fair Trade/Organic Coffee, Rural Livelihoods, and the "Agrarian Question": Southern Mexican Coffee Families in Transition. *World Development*, 39(1), pp.134-145.
- Becchetti, Leonardo, Stefano Castriota, and Pierluigi Conzo (2015). "Quantitative Analysis of the Impacts of Fair Trade." In Reynolds, L.T., and E.A. Bennett, editors. *Handbook of Research on Fair Trade*. Edward Elgar Publishing.
- Beuchelt, T.D., and M. Zeller (2012). The role of cooperative business models for the success of smallholder coffee certification in Nicaragua: A comparison of conventional, organic and Organic-Fairtrade certified cooperatives. *Renewable Agriculture And Food Systems*, 28(3), pp.195-211 (Combined w/B&Z (2011) because same data, similar conclusions.)
- Beuchelt, T.D., and M. Zeller (2011). Profits and poverty: Certification's troubled link for Nicaragua's organic and fairtrade coffee producers. *Ecological Economics*, 70(7), pp.1316-1324.
- Beuchelt, T., M. Zeller, and T. Oberthur (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.
- Bitzer, V., P. Glasbergen, and B. Arts (2013). Exploring the potential of intersectoral partnerships to improve the position of farmers in global agrifood chains: findings from the coffee sector in Peru. *Agriculture and Human Values*, 30(1), pp.5-20.
- Blackman, Allen, and Jorge E. Rivera (2010). "The evidence base for environmental and socioeconomic impacts of 'sustainable' certification."
- Chengappa, P.G., K.M. Rich, A. Muniyappa, C.G. Yadava, M.K. Ganashruthi, B.N. Pradeepa Babu, Y.C. Shubha, and M. Rich (2014). *Sustainability coffee certification in India: Perceptions and Practices*. Norwegian Institute of International Affairs.
- Chiputwa, B., D.J. Spielman, and M. Qaim (2015). Food standards, certification, and poverty among coffee farmers in Uganda. *World Development*, 66, pp.400-412.
- Chiputwa, Brian, and Matin Qaim (2016). "Sustainability standards, gender, and nutrition among smallholder farmers in Uganda." *The Journal of Development Studies* 52.9: 1241-1257.
- Coles, C (2011). Kilimanjaro and Oromia Coffee Value Chain Case Studies: Producer Benefits from Fair Trade and Free Market Channels. *NCCR North-South Dialogue*, 24.

- COSA (2013). The COSA Measuring Sustainability Report: Coffee and Cocoa in 12 Countries. Philadelphia, PA: The Committee on Sustainability Assessment.
- Dammert, A. C., and S. Mohan (2015). A survey of the economics of fair trade. *Journal of Economic Surveys*, 29(5), 855-868.
- De Janvry, Alain, Craig McIntosh, and Elisabeth Sadoulet (2015). "Fair Trade and Free Entry: Can a Disequilibrium Market Serve as a Development Tool?" *The Review of Economics and Statistics*, July 2015, 97(3): 567–573.
- Dietz, Thomas, Jennie Auffenberg, Andrea Estrella Chong, Janina Grabs, and Bernard Kilian (2018). "The Voluntary Coffee Standard Index (VOCSI): Developing a Composite Index to Assess and Compare the Strength of Mainstream Voluntary Sustainability Standards in the Global Coffee Industry." *Ecological Economics*, vol. 150 (August): pp. 72-87.
- Dietz, Thomas, Janina Grabs, and Andrea Estrella Chong (forthcoming). "The Effectiveness of Mainstreamed Voluntary Sustainability Standards – Evidence from the Honduran Coffee Sector."
- Donovan, J., and N. Poole (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 N. Poole (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.
- Dragusanu, Raluca, Daniele Giovannucci, and Nathan Nunn (2014). "The Economics of Fair Trade." *Journal of Economic Perspectives*, Volume 28, Number 3 (Summer): pp. 217–236.
- Dragusanu, Raluca, and Nathan Nunn (2014). The Impacts of Fair Trade Certification: Evidence from Coffee Producers in Costa Rica (Preliminary and Incomplete), 28 February.
- El Ouaamari, S. and H. Cochet (2014). The Role of Coffee in the Development of Southwest Ethiopia's Forests: Farmers' Strategies, Investor Speculation, and Certification Projects. *Society & Natural Resources*, 27(2), pp.200-214.
- Elder, S.D., H. Zerriffi, and P. Le Billon (2012). Effects of fair trade certification on social capital: the case of Rwandan coffee producers. *World Development*, 40(11), pp.2355-2367.
- Elliott, Kimberly Ann (2012). Is My Fair Trade Coffee Really Fair? Trends and Challenges in Fair Trade Certification. CGD Policy Paper. Washington: Center for Global Development.
- Estrella, A (forthcoming). Coffee Certifications: A Profitable Strategy for Improving the Livelihoods of Smallholder Farmers? Evidence from an Impact Evaluation in Colombia's Coffee Belt.
- Fort, R., and R. Ruben (2008). "The impact of fair trade on coffee producers in Peru." In R. Ruben (ed.), *The impact of Fair Trade*. Netherlands: Wageningen Academic Publishers. Chapter 3.
- Francesconi, G. N., and R. Ruben (2014). FairTrade's theory of change: an evaluation based on the cooperative life cycle framework and mixed methods. *Journal of Development Effectiveness*, 6(3), 268-283.

- Fraser, J., E. Fisher, and A. Arce (2014). Reframing 'crisis' in fair trade coffee production: Trajectories of agrarian change in Nicaragua. *Journal of Agrarian Change*, 14(1), 52-73.
- Grabs, Janina, Bernard Kilian, Daniel Calderón Hernández, and Thomas Dietz (2016). "Understanding Coffee Certification Dynamics: A Spatial Analysis of Voluntary Sustainability Standard Proliferation." *International Food and Agribusiness Management Review*, Volume 19 Issue 3.
- Hagggar, Jeremy, et al. (2012) "Environmental and economic costs and benefits from sustainable certification of coffee in Nicaragua." *Food Chain* 2.1: 24-41.
- Haight, Colleen (2011) *The Problem with Fair Trade Coffee*. Stanford Social Innovation Review, Summer.
- Hoebink et al. (2014). "The Impact of UTZ Certification on Smallholder Farmers in Uganda." In P. Hoebink, R. Ruben, W. Elbers, and B. van Rijsbergen, eds. *The impact of coffee certification on smallholder farmers in Kenya, Uganda and Ethiopia*. Report prepared by Centre for International Development Issues Nijmegen, Radboud University Nijmegen, for Solidaridad.
- Hoffmann, U. and F. Grothaus (2015). *Assuring Coherence between the Market-access and Livelihood Impact of Private Sustainability Standards*. United Nations Forum on Sustainability Standards (UNFSS), CH-Geneva.
- Ibnu, Muhammad, et al. (2015). "Farmer preferences for coffee certification: A conjoint analysis of the Indonesian smallholders." *Journal of Agricultural Science* 7.6: 20.
- 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).
- Jena, Pradyot Ranjan, Bezawit Beyene Chichaibelub, Till Stellmachera, Ulrike Grotea (2012). "The impact of coffee certification on small-scale producers' livelihoods: a case study from the Jimma Zone, Ethiopia." *Agricultural Economics* 43: 429-440.
- Johannessen, S., and H. Wilhite (2010). Who really benefits from Fairtrade? An analysis of value distribution in Fairtrade coffee. *Globalizations*, 7(4), pp.525-544.
- Kiemen, A., and T. Beuchelt (2010). Certification as an upgrading strategy for small-scale farmers and their cooperatives : a value chain analysis for Nicaraguan coffee. *Research in Development Economics and Policy (Discussion Paper Series)*, (2/2010), pp.570-582.
- Kirumba, E.G., and F. Pinard (2010). Determinants of farmers' compliance with coffee eco-certification standards in Mt. Kenya region. In: *Joint 3rd African Association of Agricultural Economists (AAAE) and 48th Agricultural Economists Association of South Africa (AEASA) Conference*, Cape Town, South Africa.
- Kodama, Y. (2009). The Effects of Fair Trade on Coffee Producers: A Case Study of Ethiopian Coffee Cooperatives. In: *Proceedings of the 16th International Conference of Ethiopian Studies*, Trondheim, Norway, pp. 779-792.
- K. P. M. G. Sustainability (2013). "Improving smallholder livelihoods: Effectiveness of certification in coffee, cocoa and cotton." 126-2013.
- Krupka, J. (2012). *The Fair Trade Coffee Business Model's Affect on the Small Scale Producers through the Lens of the Triple Bottom Line*. PhD. Georgia State University.
- Kuit, Michiel, Fédés van Rijn, Vu Thi Minh Tu, and Pham Van Anh (2013). *The Sustainable Coffee Conundrum: A study into the effects, cost and benefits of implementation*

- modalities of sustainable coffee production in Vietnam. Kuit Consultancy, commissioned by DE Foundation and the Dutch Ministry of Economic Affairs.
- Kuit, Michiel, and Yuca Waarts (2014). Small-scale farmers, certification schemes and private standards: Costs and benefits of certification and verification systems for small-scale producers in cocoa, coffee, cotton, fruit and vegetable sectors. Technical Centre for Agricultural and Rural Cooperation, Wageningen.
- Lazaro, E., J. Makindara, and F. Kilima (2008). Sustainability Standards and Coffee Exports From Tanzania. Copenhagen, Denmark: Danish Institute for International Studies.
- Luna, F., and P.N. Wilson (2015). An Economic Exploration of Smallholder Value Chains: Coffee Transactions in Chiapas, Mexico. *International Food and Agribusiness Management Review*, (18)3, pp. 85-106.
- Madjidi, O. (2011). Sustainable coffee certification programs and coffee cooperatives in Guatemala: A small-scale producer perspective. Masters. Royal Roads University.
- Mendez et al. (2010). Effects of Fair Trade and organic certifications on small-scale coffee farmer households in Central America and Mexico. *Renewable Agriculture and Food Systems*: 25(3); 236–251.
- 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), pp. 577-591.
- Milford, A.B. (2012). The Pro-Competitive Effect of Coffee Cooperatives in Chiapas, Mexico. *Journal of Agricultural & Food Industrial Organization*, 10(1).
- Minten, B., M. Dereje, E. Engeda, and S. Tamru (2015). Who benefits from the rapidly increasing Voluntary Sustainability Standards? Evidence from Fairtrade and Organic certified coffee in Ethiopia. IFPRI Working Paper 71.
- Mitiku, Fikadu, et al. (2015). "Do Private Sustainability Standards Contribute to Poverty Alleviation? A Comparison of Different Coffee Certification Schemes in Ethiopia." *Bioeconomics Working Paper Series: WP 2015/03*. [Updated version published in *Sustainability*, (2017, no. 9).
- Neilson, J., and R. Toth (2016). Baseline report: evaluation of the early impacts of sustainability standards on smallholder coffee farmers in Lampung and South Sumatra, Indonesia. Sydney: University of Sydney.
- Neilson, Jeff, and Bill Pritchard (2010). "Fairness and ethicality in their place: the regional dynamics of fair trade and ethical sourcing agendas in the plantation districts of South India." *Environment and Planning* 42.8: 1833-1851.
- Oya C, F. Schaefer, D. Skalidou, C. McCosker, L. Langer (2017). Effects of certification schemes for agricultural production on socio-economic outcomes in low- and middle-income countries: a systematic review. *Campbell Systematic Reviews* 2017:3
- Pinto, Luís Fernando Guedes, et al. (2014). "Group certification supports an increase in the diversity of sustainable agriculture network–rainforest alliance certified coffee producers in Brazil." *Ecological Economics* 107: 59-64.
- Pongvinyoo, P. and M. Yamao (2014). Coffee farmers attitudes toward the 4C process in Chumphon province, Southern Thailand. *Journal of Agricultural Extension and Rural Development*, 6(8), pp.249-258.
- Potts, Jason, Matthew Lynch, Ann Wilkings, Gabriel Huppe, Maxine Cunningham, and Vivek Voora (2014). *The State of Sustainability Initiatives Review 2014: Standards and*

- the Green Economy. Winnipeg, Manitoba, and London: International Institute for Sustainable Development and International Institute for Environment and Development.
- Powae, W.I. (2009). Fair trade coffee supply chains in the highlands of Papua New Guinea: do they give higher returns to smallholders? Masters. Lincoln University.
- Riisgaard, L., G. Michuki, P. Gibbon, and S. Bolwig (2009). The performance of voluntary standard schemes from the perspective of small producers in East-Africa. Copenhagen, Denmark: Danish Institute for International Studies.
- Ruben, R., et al. (2014). "The Impact of Coffee Certification on Smallholder Farmers in Kenya." In Hoebink, P., Ruben, R., Elbers, W., and van Rijsbergen, B., ed. The impact of coffee certification on smallholder farmers in Kenya, Uganda and Ethiopia. Report prepared by Centre for International Development Issues Nijmegen, Radboud University Nijmegen, for Solidaridad.
- Ruben, R., and R. Fort (2012). The impact of fair trade certification for coffee farmers in Peru. *World Development*, 40(3), pp.570-582.
- Ruben, R., and G. Zúñiga-Arias (2011). How standards compete: comparative impact of coffee certification schemes in Northern Nicaragua. *Supply Chain Management: An International Journal*, 16(2), pp.98-109. [unclear risk of bias but many flags]
- Ruben, Rued, Ricardo Fort & Guillermo Zúñiga-Arias (2009): Measuring the impact of fair trade on development, *Development in Practice*, 19:6, 777-788.
- Rueda, X., and E.F. Lambin (2013). Responding to Globalization: Impacts of Certification on Colombian Small-Scale Coffee Growers. *Ecology and Society*, 18(3), 21.
- Sáenz Segura, F., and G. Zúñiga-Arias (2008). "Assessment of the effect of Fair Trade on smallholder producers in Costa Rica: A comparative study in the coffee sector." In R. Ruben (ed.), *The impact of Fair Trade*. Netherlands: Wageningen Academic Publishers. Chapter 5.
- Smith, Sally (2015). "Fair Trade and Women's Empowerment." In Reynolds, L.T., and E.A. Bennett, editors. *Handbook of Research on Fair Trade*. Edward Elgar Publishing.
- Snider, Anna, Eva Kraus, Nicole Sibelet, Aske Skovmand Bosselmann and Guy Faure (2016). Influence of voluntary coffee certifications on cooperatives' advisory services and agricultural practices of smallholder farmers in Costa Rica, *The Journal of Agricultural Education and Extension*.
- Staib, P. W. (2012). *Coffee and the countryside: Small farmers and sustainable development in Las Segovias de Nicaragua*. The University of New Mexico.
- Terstappen, V., L. Hanson, and D. McLaughlin (2013). Gender, health, labor, and inequities: a review of the fair and alternative trade literature. *Agriculture and Human Values*, 30(1), pp.21-39.
- University of London, SOAS (2014). *Fairtrade, Employment and Poverty Reduction in Ethiopia and Uganda: The Final Report to DFID*, April.
- Valkila, J. (2014). Do Fair Trade pricing policies reduce inequalities in coffee production and trade? *Development Policy Review*, 32(4), pp.475-493.
- Valkila, J. (2009). Fair Trade organic coffee production in Nicaragua: Sustainable development or a poverty trap? *Ecological Economics*, 68, pp.3018-3025.

- Valkila, J., and A. Nygren (2008). Impacts of Fair Trade-certification on Coffee Farmers, Cooperatives and Laborers, in Nicaragua. *Agriculture and Human Values*, 27(3), pp.321-333.
- Valkila, J., P. Haaparanta, and N. Niemi (2010). Empowering Coffee Traders? The Coffee Value Chain from Nicaraguan Fair Trade Farmers to Finnish Consumers. *Journal of Business Ethics*, 97(2), pp.257-270.
- van Rijsbergen, B., W. Elbers, R. Ruben, and S.N. Njuguna (2016). The ambivalent impact of coffee certification on farmers' welfare: a matched panel approach for cooperatives in central Kenya. *World Development*, 77, pp.277-292.
- Weber, J.G. (2011). How Much More Do Growers Receive for Fair Trade-Organic Coffee? *Food Policy*, 36(5), pp.678-685.
- Wilson, B.R. (2010). Indebted to Fair Trade? Coffee and crisis in Nicaragua. *Geoforum*, 41(1), pp.84-92.

Selected older studies

- Arnould, E.J., A. Plastina, and D. Ball (2006). Market Disintermediation and Producer Value Capture: The Case of Fair Trade Coffee in Nicaragua, Peru and Guatemala. In: *Product and Market Development for Subsistence Marketplaces: Consumption and Entrepreneurship Beyond Literacy and Resource Barriers*. Chicago: University of Illinois at Chicago.
- Calo, M., and T. Wise (2005). "Revaluing Peasant Coffee Production: Organic and Fair Trade Markets in Mexico. Globalization and Sustainable Development Program, Tufts University.
- Consumers International and the International Institute for Environment and Development (2005). *From bean to cup: How consumer choice impacts upon coffee producers and the environment*. December.
- Fitter, R., and R. Kaplinsky (2001). "Who Gains from Product Rents as the Coffee Market Becomes More Differentiated? A Value Chain Analysis." *IDS Bulletin*, vol. 32, no. 3: 69-82.
- Fridell, Gavin (2007). *Fair Trade Coffee: The Prospects and Pitfalls of Market-Driven Social Justice*. Toronto: University of Toronto Press.
- Giovannucci, Daniele, and Jason Potts, with B. Killian, C. Wunderlich, G. Soto, S. Schuller, F. Pinard, K. Schroeder, I. Vagneron (2008). *Seeking Sustainability: COSA Preliminary Analysis of Sustainability Initiatives in the Coffee Sector*. Winnipeg, Manitoba: Committee on Sustainability Assessment.
- Imhof, Sandra, and Andrew Lee (2007). Assessing the Potential of Fair Trade for Poverty Reduction and Conflict Prevention: A Case Study of Bolivian Coffee Producers (short version).
- Jaffee, Daniel (2007). *Brewing Justice: Fair Trade Coffee, Sustainability, and Survival*. Berkeley, CA: University of California Press.
- Jones, S. and B. Bayley (2000). Fair Trade: Overview, Impact, Challenge: Study to Inform DFID's Support to Fair Trade. Oxford, Oxford Policy Management.
- Kilian B, L. Pratt, C. Jones, and A. Villalobos (2004). Can the Private Sector be Competitive and Contribute to Development through Sustainable Agricultural Business? A Case

- Study of Coffee in Latin America. *International Food and Agribusiness Management Review* Volume 7, Issue 3.
- Mayoux, Linda (2004). "Impact Assessment of Fair Trade and Ethical Enterprise Development." Enterprise Development Impact Assessment Information Service, Manchester.
- Murray, Douglas L., Laura T. Reynolds & Peter L. Taylor (2006). The future of Fair Trade coffee: dilemmas facing Latin America's small-scale producers, *Development in Practice*, 16:02, 179-192.
- Ronchi, L. (2002). The impact of Fair Trade on producers and their organizations: A case study with Coocafé in Costa Rica. Policy Research Unit. Sussex: University of Sussex.
- Pirotte, Gautier, Geoffrey Pleyers and Marc Poncelet (2006). Fair-trade coffee in Nicaragua and Tanzania: a comparison, *Development in Practice*, 16:5, 441-451.
- Vanderhoff Boersma, F. (2002). "Poverty Alleviation through Participation in Fair Trade Coffee Networks: The Case of UCIRI." Colorado State University.

Tables and Figures

Table 1. Ranking of initiatives on coverage of key issues, by dimension*
(for Potts, percent coverage of indicators in parentheses; for Dietz, score in parentheses)

Initiative	Economic		Social (worker, human rights, gender)		Environmental	
	Potts	Dietz	Potts	Dietz	Potts	Dietz
4C	4 (8)	4 (12.3)	4 (33)	4 (40.5)	4 (35)	4 (24.8)
Fairtrade	1 (68)	1 (59.1)	2 (73)	3 (62.7)	2 (60)	3 (35.1)
Rainforest Alliance	3 (20)	3 (24.3)	1 (84)	2 (78.2)	1 (74)	1 (67.0)
UTZ Certified	2 (40)	2 (51.7)	3 (58)	1 (86.8)	3 (43)	2 (65.2)

* Adjusted to reflect rank of just these four within ranking of many more initiatives.

Sources: Potts et al. (2014, pp. 71, 73, 76; Dietz et al. (2018).

Table 2. Audit requirements by initiative

Initiative	Audit frequency	Sample selection	Notification
4C/CAS	Every 3 years; self-assessment annually	50% of square root of number of “business partners,” based on risk assessment; a portion must be resampled in next audit to ensure continuous improvement Workers “if available,” with “all efforts” to ensure confidentiality; verifier to determine number	Audit is scheduled with group management but sampled partners only notified 2 days in advance Unannounced visits “as applicable,” with no more than 2 days notice
FT	Full audits every 3 years, 1 surveillance audit in interim, except high risk groups annually; small licensees, 6-year cycle with 1 full, 1 surveillance audit (desk only for latter)	Different minimum numbers set for members and workers at different types of organizations, based on size. For first grade producer organizations, the minimums tend to be just above the square root of the upper end of range. Beyond 1,000 members, the minimum is 40, which would become less than the square root at 1,600. Auditors may choose to interview more. Auditors should aim to interview workers at each farm visited, as well as interviewing at processing plants.	Regular audits arranged in advance; unannounced involve no notice and can bring suspension if refused
RA	Full audits every 3 years with 2 surveillance audits in interim for critical criteria, evidence of continuous improvement	Square root of number of producers in group, but up to 1.4 times that for high risk, or half that for surveillance audits of “high-performing clients” (note still working to incorporate risk management into process as of December 2015)	Full audits scheduled in advance; surveillance audits with no more than 2 days notice; unannounced audits can occur in response to complaints, reported incidents, other “substantial information”
UTZ	Annual	Square root of number of plots (individually certified farm) or producers in group; each site for multi-site estates; randomly selected plus high risk; different sample each time Worker interviews should be anonymous and w/out supervisors, but need not be off premises	Surprise audits < 72 hours notice, designed to check auditors as well as producers; Certification bodies with > 10 certificates issued/year must do shadow audits, though they decide sample and code principles to audit, samples chosen based on risk assessment or complaints received (if any)

Table 3. Key indicators of VSS market status (2015 unless otherwise noted)

	Hectares (‘000)	Volume (‘000 MT)	Volume sold as cert. (‘000 MT)	Number of certificates	Number of producers	Major sources, by area certified
4C/GCP	1,594	2,629	610 (23%)	321	505,000	50% Brazil; 25% Colombia
Fairtrade (2014)	1,297 (2015)	549	151 (28%)	445 POs	813,000	16% each Colombia, Ethiopia
Rainforest Alliance (2014)	405 (2015)	457	168 (37%)		194,000 (158,000 ≤ 2 ha)	20% Brazil; 40% other LatAm
Utz Certified (2016) (global avg. premium of \$0.03, 2014, down from \$0.04, 2012)	567	870 (estimated potential) (40% Brazil estates)	321 (37%) (78% LatAm)	563 (2014) (farmers, estates, producer groups)	225,505 farmers 919 ind. farms, estates 296,208 workers	59% Africa (farmers) 85% LatAm (estates) 52% LatAm (workers)

NB: Does not reflect potential multiple certification and therefore overstates total acreage and volume certified by an unknown amount.

Table 4. Findings on impact from previous reviews of studies with counterfactuals

Indicator assessed	Positive	Neutral or mixed	Negative
Price			
ITC	3	1	
K&W	8	4	0
This review	6		1
Output effects			
ITC	2	1	1
K&W	2	5	1
This review	2	3	2
Quality			
ITC		1	
K&W	2	1	
Income			
ITC	1	2	
K&W	2	8	1
This review	3	3	1
Business opportunities			
ITC	2	1	
K&W (market access)	1	2	0
Producer livelihoods			
ITC	4	1	
K&W (not reported)			
Working conditions			
ITC (none)			
K&W (child labor)	0	2	0
This review	2	3	
Environment			
ITC	1		1
K&W	2	2	
This review	2	3	

NB: ITC (2011) review includes 5 coffee studies, all FT. Kuit and Waarts (2014) (K&W) include 14 coffee studies, up to 5 FT, at least 1 4C, 2 each Rainforest Alliance and UtTZ, also organic, C.A.F.E. There is no distinction among the initiatives in counting impacts. The studies added for this review are Akoyi and Maertens (2016), which has case studies of FT/organic (double certified) and UTZ/RA/4C (triple certified) in Uganda, and Dietz et al. (forthcoming), which has case studies of all four initiatives assessed in this paper, as well as FT/org (double certified).

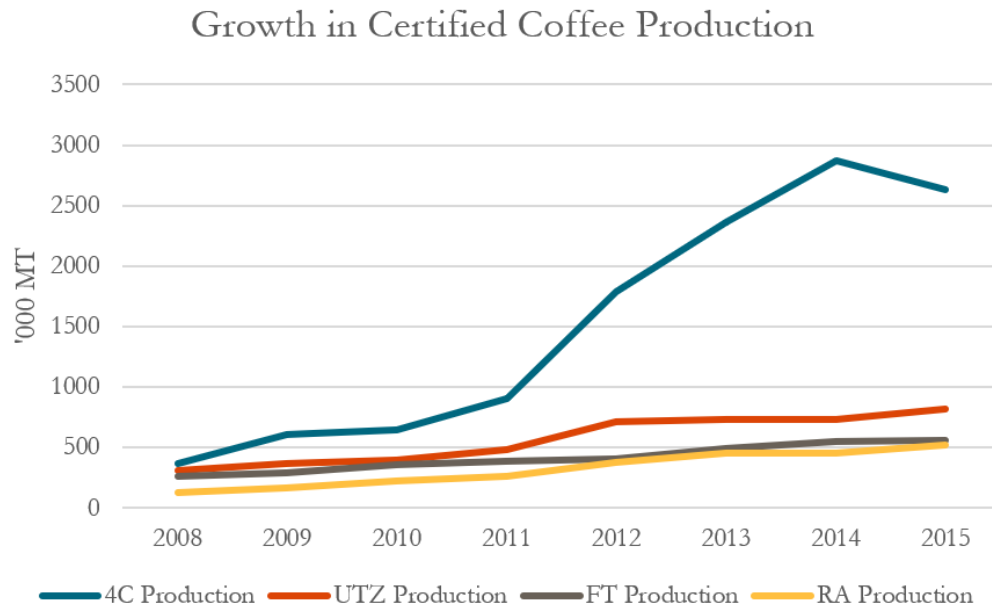
Table 5. Summary of quantitative effects from Campbell Review

Effect	Effect size, all products	# of studies/ # with high risk of bias	Coffee-specific results where available (standardized mean difference/risk of bias)
Price	+14%, statistically significant	4/1	-0.03 (moderate) FT, Peru 0.19 (moderate) FT, Mexico 0.42 (moderate) Various, Ethiopia
Yield	-20%, not statistically significant	5/2	-2.20 (high) FT, Ethiopia -0.32 (moderate) FT, Peru 0.19 (low) FT, Kenya
Income from certified production	+11%, statistically significant	10/4	-0.17 (moderate) FT, Peru -0.02 (moderate) FT, Uganda 0.25 (low) FT, Kenya 0.80 (moderate) Utz, Uganda
Wages	-13%, statistically significant	8/2	-0.39 (moderate) FT, Ethiopia -0.26 (moderate) FT, Uganda 0.01 (high) FT, Costa Rica
Household income (includes net and gross)	+6%, not statistically significant	8/4	-0.28 (moderate) FT, Peru -0.09 (high) FT, Ethiopia 0.48 (moderate) FT, Uganda
Assets/wealth	+3%, not statistically significant	Only 2	None
Health (illness incidence)	-7%, not statistically significant	Only 2	None
Schooling	+6%, statistically significant	5/3	-0.01 (high) FT, Costa Rica 0.36 (moderate) Various, Ethiopia

Note substantial heterogeneity and low robustness due to small number of studies; note also that only the figures reported for coffee in bold type are statistically significant.

Source: Compiled from results reported in Oya et al. (2017).

Figure 1.



Source: Potts et al. (2014).