Evaluating the Sustainability Performance of Typical Conventional and Certified Coffee Production Systems in Brazil and Ethiopia Based on Expert Judgements

Eva Winter*, Silvia M. R. R. Marton, Lukas Baumgart, Michael Curran, Matthias Stolze and Christian Schader

Research Institute of Organic Agriculture (FiBL), Frick, Switzerland

Increasing consumer awareness on sustainability issues has led to the growing adoption of voluntary sustainability standards in agriculture. This study assesses the sustainability performance of typical conventional and certified coffee production systems in Brazil and Ethiopia based on expert judgements. We apply the SMART-Farm Tool, which represents an operationalization of the SAFA (Sustainability Assessment of Food and Agriculture systems) framework of FAO. Data were collected through expert interviews and uncertainties were estimated using Monte-Carlo simulations. A higher sustainability performance of the certified systems was observed regarding product information (+37%) and transparency (+39%) in Ethiopia. In Brazil, the certified system showed a higher overall sustainability performance compared to the conventional system in the environmental dimension and in some social and governance aspects, e.g., gender equality (+49%) and public health (+36%). Geographical or political conditions and farm type also had a strong influence on the observed sustainability performance. Typical smallholder production systems in Ethiopian coffee production performed comparable in the environmental dimension since all were low-input systems due to economic constraints. The conventional Brazilian system showed a better performance concerning employment relations (+14%) and profitability (+13%), as compared to the certified Brazilian systems, because larger farms were more likely to employ permanent staff and benefit from economies of scale.

Keywords: sustainability assessment, SMART-Farm Tool, RRID:SCR_018197, organic, coffee, fair trade

INTRODUCTION

In recent decades, consumers have become more aware of issues in coffee production such as socio-economic challenges of smallholder farmers and environmental degradation caused by common practices in conventional agriculture (Pierrot et al., 2011). This has led to increased demand for sustainability certification schemes, such as organic and Fairtrade (FLO) voluntary sustainability standards (VSS). Retailers use these schemes to convince consumers and other partners in the value chain (e.g., governments and non-governmental institutions) of the particular sustainability benefits of their products (Pierrot et al., 2011). Consequently, the questions arise as to

the validity of sustainability claims around these VSS in coffee production relative to conventional practices. Different VSS have varying requirements for certification, generally regarding social, economic, or environmental aspects of the farming system. These include minimum farm gate prices, trainings for farmers and defined procedures for handling pesticides. Verification of compliance is generally assured by a public or private certification body, participation is voluntary, and requirements can go beyond legal requirements (Lernoud et al., 2016). Organic certification stands for a system-oriented approach, which strives to operate as close to a closed cycle as possible. Key elements are the prohibition of easily-soluble mineral fertilizers, chemical synthetic pesticides, and genetically modified organisms (Piras, 2011). The principle objectives of Fairtrade certification include fair minimum producer prices, a Fairtrade Premium, pre-financing of inputs, long-term trading relationships, and regulations to ensure socially and economically fair and environmentally responsible production and trading conditions (FLO, 2017). We chose these two VSS, because organic and Fairtrade certification schemes are widespread and well known (Lernoud et al., 2016). Furthermore, both schemes require full compliance as a pre-condition of certification (European Commission, 2007; FLO, 2017). This makes evaluating certified and non-certified farming systems possible.

In order to evaluate systems in a country with a very low human development index and an emerging country, Ethiopia and Brazil were chosen as case studies. Amongst coffee producing nations, Brazil and Ethiopia represent long-established, global Arabica producers. Brazil is by far the world's largest producer of green coffee, with a coffee cultivation history going back to the late 18th century (Boddey et al., 2003; FAOSTAT, 2016). Environmental pollution and negative health effects on farmers and surrounding inhabitants through the over-use of chemical pesticides are major difficulties in Brazilian coffee production (Boddey et al., 2003; Carvalho, 2006). Ethiopia is the historical country of origin of Arabica coffee and the biggest Arabica coffee producing country in Africa. Additionally, coffee is one of the economically most important commodities produced and exported by Ethiopia (FAOSTAT, 2016; Tefera, 2016). Difficulties regarding social and economic sustainability in Ethiopian coffee production are frequently mentioned in the literature (Jena et al., 2012; Minten et al., 2015). With regard to certification, Ethiopia is the largest African exporter of organic and Fairtrade certified coffee. In the main coffee producing regions of both countries double certification of these VSS is wide-spread (Minten et al., 2015; Pedini, 2016).

Several assessments have been conducted in the two countries to investigate how different VSS applied in coffee affect sustainability. Looking at Ethiopia, results indicate that organic and Fairtrade certifications do not have a significant positive impact on coffee farmers' livelihoods, as the price premium transmission down to the farm level is low (Jena et al., 2012; Minten et al., 2015; Tefera, 2015; Abdissa et al., 2017). One study showed that Rainforest Alliance (RFA) certified Ethiopian coffee farmers profited significantly from their certification because the value chain established in this scheme was very short and

farmers could reap most of the price premium themselves (Abdissa et al., 2017). Another study assessed differences in livelihood indicators of farms that were certified with a single scheme vs. farms that were certified with more than one certification scheme in Ethiopia. No significant differences could be identified regarding livelihood indicators such as "access to credit" and "higher prices," when comparing Fairtrade and Fairtrade/organic double certified farms. Only between single [FLO and triple (FLO/organic/UTZ (formerly UTZ Kapeh, which means "good coffee" in the Mayan language)] certification, significant differences in the applied indicators were found (Woubie et al., 2015). A few studies have been conducted attempting to assess sustainability in Brazilian coffee production, mainly with regard to only one or two sustainability dimensions. One study reported a positive correlation between RFA certification and biodiversity (Hardt et al., 2015). Another assessed the socio-economic performance of organic coffee producers, finding that family farms perform well in this respect and that larger farms have difficulties to achieve socio-economic resilience under organic management (Wegner et al., 2013). However, there is a lack of data on the coffee sector that covers all dimensions of sustainability according to an international and established sustainability framework (Pierrot et al., 2011). This paper addresses this gap to deliver new information on coffee production sustainability in the two mentioned case studies.

Many different sustainability evaluation methods have been developed in recent years in the food sector. According to The Food and Agriculture Organization of the United Nations Organization (FAO), 106 countries have their own National Sustainable Development Strategies and at least 170 voluntary sustainability standards exist world-wide in the food and agriculture industry (FAO, 2016). The FAO developed the "Guidelines for Sustainability Assessment of Food and Agriculture systems" (SAFA) with the aim to create a holistic global framework for assessing sustainability along food and agricultural value chains based on an international reference system. The SAFA Guidelines include four dimensions, 21 themes and 58 subthemes, each with defined sustainability targets. With this framework, the FAO established a globally applicable and comprehensive assessment framework of sustainability. **Figure 1** provides an overview on the structure of the Guidelines. In this study, we used the SMART-Farm Tool, which stands for **S**ustainability **M**onitoring and **A**ssessment **R**outine. It is a multi-criteria assessment approach including an impact matrix with a set of 327 science-based indicators and 1,769 relations between indicators within the 58 SAFA subthemes (Schader et al., 2016). It also takes account of both primary production at farm level as well as upstream and downstream value chain (such as purchased inputs and marketing channels) into account (Schader et al., 2016). The method has been applied to assess sustainability of individual farms in both high and low-income countries (Schader et al., 2016, 2019), including coffee production under different VSS in Uganda (Ssebunya et al., 2019). Based on this aforementioned study, certification improves subtheme goal achievement in Uganda through an increased group organization and collective capacities.
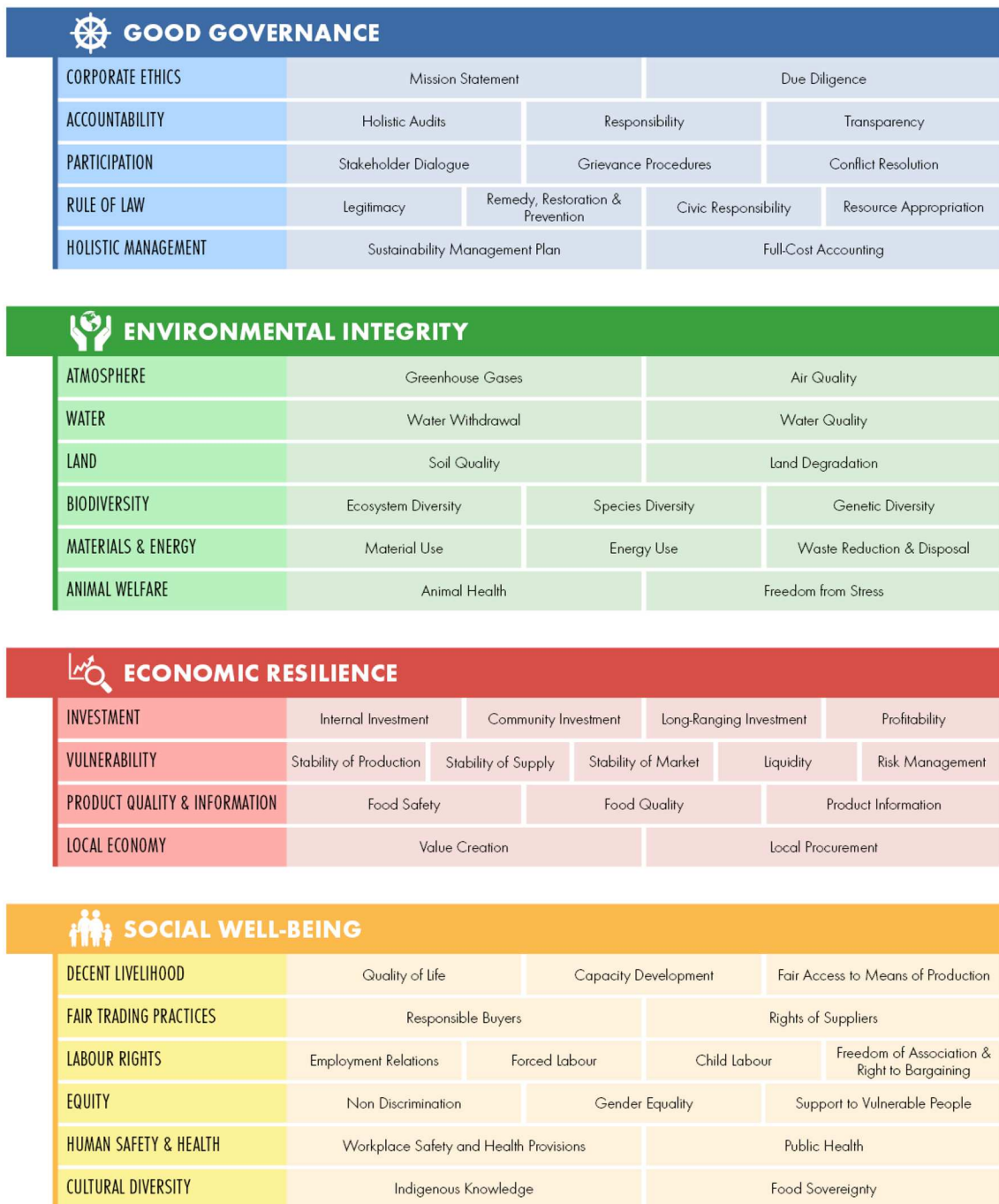
GOOD GOVERNANCE

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

ENVIRONMENTAL INTEGRITY

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

ECONOMIC RESILIENCE

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

SOCIAL WELL-BEING

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

FIGURE 1 | Dimensions, themes, and subthemes of the SAFA Guidelines (FAO, 2014).

However, a comprehensive sustainability assessment has not yet been conducted that makes evaluating coffee production systems across countries and across farming systems using a consistent approach. This paper achieved this with an evaluation of Ethiopian and Brazilian systems using a modeling approach driven by expert judgements. The concept of a typical or average farm (Further explanations in Typical farm theory) as assessment object is introduced to represent subsets of the sector as a whole. Through the application of the SMART-Farm Tool, this study answers the following research questions:

- How do typical coffee production systems in Ethiopia and Brazil perform in the different dimensions of sustainability?

- Where can similarities and differences between conventional and certified systems be observed in the two country cases?

MATERIALS AND METHODS

The procedure for selecting typical production systems is described in the following two subsections. In the following subsection, the SMART-Farm Tool and how it was applied in this study is introduced. Further, we explain the data collection process. Finally, in the last subsection the applied uncertainty analysis is laid out.

Typical Farm Theory

We define a typical farm as a farm type that represents the mode of the distribution of farms according to defined classification criteria. The farm depicting the mode is the farm that can be found the most often on the ground and thus differs from an average farm that depicts the mean of all farms and which rarely exists on the ground. We deem the definition of a typical farm through the mode as more meaningful in the context of sustainability assessments and have included a more extensive justification of this topic in the discussion of methods section. Classification criteria are the defined geographical area, the relevant farm enterprises, and resource endowments (Feuz and Skold, 1992). Further relevant classification criteria for this study are defined and discussed in the following subsection. The concept of typical farms has been used in many studies as a basis for e.g. policy assessments (Häring, 2003b; Thünen Institute, 2016; Reidsma et al., 2019).

Definition of Typical Coffee Production Systems in Selected Regions in Ethiopia and Brazil

The authors defined typical organic and non-certified coffee production systems and the system boundaries of Ethiopian and Brazilian coffee production according to the following classification criteria. The geographical areas of interest where the typical systems are situated in the two countries were defined as the sub-national region(s) where most coffee is produced. This was first determined through different sources of literature (sources are indicated in subsection Typical farms in Ethiopia and Brazil), and then supplemented and verified by experts (the definition and selection of experts can be found in subsection Selection of experts). Furthermore, we identified typical characteristics regarding technology constraints, resource access and management and validated these as well through literature and expert interviews. The following decision criteria were chosen to identify an organic and a non-certified coffee production system:

- What is the typical farm size, labor use, and availability? Which coffee species are grown?
- What are the specifications of the area?
- What is the typical coffee yield?
- Which other crops are grown in the typical coffee production systems and what are the typical agricultural practices?
- Is livestock kept in the typical coffee production system?

- Is the typical organic system also certified with another VSS?

SMART-Farm Tool

The SMART-Farm Tool was used as follows: all SMART indicators relevant to the respective typical systems were rated with a performance score reflecting the degree of goal achievement. This was conducted with the help of scientific literature and expert judgements (Further explanations in Selection of experts). Each indicator has an assigned weight according to its importance for the subtheme. The weights were defined by a panel of 67 experts from 21 countries using a Delphi process (Schader et al., 2019). Most indicators are relevant for several subthemes with varying importance. Weights are depicted as percentages and can be either positive or negative. The indicator achievements were assessed on different qualitative (binary and ordinal) and quantitative (numerical and percentages) benchmarks. These were then translated into percentages. For example, the indicator "Does the farm have adequate savings to cater for its cash needs?" will be answered in a qualitative way (No, Partly, Yes) which is then translated into a percentage rating (0, 50, 100%). A quantitative indicator is, e.g., "What proportion of the agricultural area does not receive synthetic chemical fungicide applications?" which is directly rated in percent of agricultural area.

For calculation of the degree of goal achievement in percent of the subthemes, the achievements of the indicators were multiplied by the weights, summed up, and divided by the sum of all possible achievements using the following formula:

$$DGA_{ix} = \frac{\sum_{n=1}^{N} (IM_{ni} * IS_{nx})}{\sum_{n=1}^{N} (IM_{ni} * IS_{max})}$$

x: index of farms
i: index of subthemes
n: index of indicators
DGA: Degree of goal achievement
IM: Impact of an indicator on a sub theme
IS: Performance of a farm with respect to indicator n

Results are categorized as follows: 0–20% insufficient, 20–40% limited, 40–60% moderate, 60–80% good, and 80–100% best sustainability performance (FAO, 2014).

Selection of Experts

According to Mieg and Näf (2005), an expert is a specialist in a certain area of knowledge with several years of experience, his/her knowledge not being transferable to another area and not being predominantly dependent on the expert's personal skills such as intelligence or memory. In this study, the definition of an expert was further clarified as either an advisor or researcher with a sufficient overview of the heterogeneity of farms in the respective country or production area to be able to rate the SMART indicators. Experts were identified through relevant scientific literature, i.e., authors of scientific publications concerning coffee production. Extension services or development agencies working with coffee producers were also included. Experts were further asked to identify other experts in their respective field of expertise (snowball sampling). This is a crucial

part of the methodology as through this one gets first-hand information on who belongs to the circle of experts in the respective area and minimizes the chances of leaving relevant interview partners out (Bogner et al., 2014).

The section of the study that involved human participants was performed in accordance with all relevant institutional and national ethical guidelines. Approval by an ethics committee was not required in accordance with Swiss law. Informed consent was obtained from respondents in accordance with section 32 of REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation).

Data Collection Through Literature Review and Expert Interviews

A literature review of scientific and gray literature, as well as governmental reports and databases was carried out to get an overview on existing data relating to SMART indicators. From this, most information on the typical production systems and some SMART indicator ratings could be obtained. All information not obtained in this way was collected through expert interviews. In total, 26 experts were interviewed. For Ethiopia, one livestock expert, seven natural scientists, one social scientist, two economists, one coffee exporter, and three agricultural advisors were interviewed. In the case of Brazil, six agronomists, two economists, and three agricultural advisors were interviewed. Each expert only rated the indicators and farm types in his or her field of expertise.

To assess the overall sustainability of the chosen typical production systems, the SMART-Farm Tool Version 4.0 was used. As typical production systems were assessed with the help of expert interviews and not through individual farm assessments, the ratings of indicators were in some cases depicted as distributions and not as precise values. This method accounted for variations within the defined typical production systems. For example, an expert might find that the proportion of arable land devoted to legumes of a coffee farmer lies realistically between 10 and 30%, all values in this range having the same probability. The underlying distribution of this rating is thus uniform. Additionally, in some cases indicators were rated by more than one expert, which can also result in an indicator rating depicted as a distribution, such as a discrete distribution if one expert rates an indicator with 40% and another with 50%.

Uncertainty Analysis

In addition to the uncertainty ranges provided by the expert scores (termed "basic uncertainty"), uncertainty of the underlying scores themselves was also estimated (termed "data uncertainty"). This uncertainty was quantified and analyzed for each indicator score. To do this, we evaluated the quality of the expert ratings or rating through literature and defined an uncertainty distribution for each rating. The uncertainty distribution parameters were based on a pedigree matrix approach. The term pedigree matrix was used here as the data quality indicators describe the source of the information, like

a genealogical table documenting the pedigree of a person. It comprises five independent criteria: "reliability," "completeness," "temporal correlation," "geographic correlation," and "further technological correlation"—each divided into five quality levels that add up to scores from one to five. For each score, a normal uncertainty distribution is assigned with the mean of zero and a variance based on expert judgement. This distribution can then be added to the basic uncertainty of each indicator (Weidema et al., 2013). As we refer in SMART to a scale from zero to 100%, the normal distributions were truncated at these boundaries. In this study, only the criteria "reliability," "temporal correlation," and "geographical correlation" and within these, only a selection of quality levels was of importance, hence the non-relevant criteria were omitted (see **Table 1**). We did this to avoid an unnecessarily high variation.

The distribution of the degree of goal achievement of each subtheme was calculated with the help of Monte-Carlo uncertainty analyses (Rubinstein and Kroese, 2011) with 1000 iterations using the @RISK excel add-in to determine the Monte-Carlo uncertainty distributions. This procedure allowed us to see all possible outcomes of a scenario including the probability of their occurrence (Palisade, 2016). Unless otherwise specified, only cases where a difference in favor of one scenario is seen in 950 of 1000 simulation runs (i.e., $p < 0.05$) are mentioned in section Results and discussion.

RESULTS AND DISCUSSION
Typical Farms in Ethiopia and Brazil

In the case of Ethiopia, four types of coffee production systems were distinguished: forest, semi-forest, garden, and plantation coffee production systems (Tefera and Tefera, 2013). The semi-forest and garden coffee cropping systems are the most relevant in Ethiopia with 50 and 40 percent of the overall production, respectively (Gole, 2015). As they are also the most prevalent under organic certification with few differences to the conventional systems, only these two mentioned systems were hence considered here and described in detail (see **Table 2**).

In Brazil, the State of Minas Gerais accounts for about 50% of the country's coffee production (Barbosa et al., 2012). Because of this, we chose it as an exemplary study area. In Minas Gerais, coffee from the Cerrado region accounts for 30% of green coffee production and Montanha coffee for around 70% (Vilela and Rufino, 2010). The Montanha farms can be further divided into three groups: small (<10 hectares); medium (10–50 hectares); and large producers (>50 hectares). They are situated in the Minas Gerais highlands. This makes mechanization and intensification difficult. Nevertheless, the large producers in particular are highly mechanized insofar as is possible in the respective terrain. The biggest farm type is likely to be the one with the highest production share [this assumption is based on numbers regarding the share of area each farm type manages according to Vilela and Rufino (2010)]. As a second model system, we assessed a typical organic farm in Minas Gerais. Personal correspondence indicated this would be represented by a Fairtrade and organic certified Montanha small producer

Quality Score (σ)	1 ($\sigma = 0$)	2 ($\sigma = 0.019$)	4 ($\sigma = 0.089$)	5 ($\sigma = 0.2$)
Reliability	Verified data based on measurements	N/A	Qualified estimate	N/A
Temporal correlation	Less than 3 years of difference to the time period of the dataset	Less than 6 years of difference to the time period of the dataset	N/A	Age of data unknown or more than 15 years of difference to the time period of the dataset
Geographical correlation	Data from area under study	Average data from larger area in which the area under study is included	N/A	N/A

(Pedini, 2016). The two chosen systems are defined in further detail in **Table 2**.

When these definitions of typical production systems were applied to the SMART-Farm Tool, the number of relevant indicators per production system was the following: Ethiopian garden systems: 227, Ethiopian semi-forest systems: 241, Brazilian conventional system: 204, Brazilian certified system: 174.

Sustainability Performance in Dimension "Environmental Integrity" at the Subtheme Level

In this and the following three subsections, the degrees of goal achievement of the subthemes are presented. These were calculated with Monte-Carlo simulations based on the expert judgements of single indicators. The goal achievement of a subtheme is always set in relation to the included indicators. The values described are the means of the degrees of goal achievement in percent in the subthemes and the error bars show the standard deviations of the Monte-Carlo simulations. To give an overview of the results, the very low and high scoring subthemes are presented here and the greatest differences and similarities between the described typical systems are highlighted and discussed. A high score indicates good performance, whereas a low score indicates unsatisfactory performance. Additional discussion of subthemes can be found in **Appendix I**. A detailed description of each subtheme and its goals can be found in **Appendix II**. Lastly, the means of indicator ratings can be found in **Appendix III**. In the following, the indicator identity number (ID) is mentioned in brackets if an indicator rating is described. If no literature source is indicated, the statement is based on expert judgement.

Typical Ethiopian Systems

In this subsection, firstly, the overall performance of the dimension of the different systems is presented. Furthermore, the highest and lowest scores are discussed. Literature suggests that certification can have positive effects on environmental outcomes (Hardt et al., 2015; Vanderhaegen et al., 2018). In several countries, organic coffee certification could reduce the use of chemical input and increased adoption of some environmentally friendly management practices, such as increasing tree cover and habitat conservation (Blackman and Naranjo, 2010; Jurjonas

et al., 2016; Giuliani et al., 2017). However, in our assessment, overall, a moderate to good performance was observed in **Figure 2** of all four Ethiopian systems, not only for the certified systems. The four assessed Ethiopian production systems are diversified and mostly extensively managed with hardly any use of external inputs that could cause contamination, regardless of certification status (Tefera, 2015). Literature describes Ethiopian smallholder coffee as around 95% organically managed even though only a small part is formally certified (Tefera, 2015). The use of synthetic pesticides and fertilizers is reported only in very few cases (Jena et al., 2012; Minten et al., 2015). Fairtrade only has minor requirements regarding the environmental dimension and thus also does not induce a better performance for the certified systems (FLO, 2017).

In this paragraph, some exemplary subthemes are further explored: For example, some of the major driving forces (impact of indicator on subtheme outcome over 50%) behind the good performance in all four systems of the subtheme "Genetic Diversity" are the following SMART-Farm Tool indicators: No use of synthetic insecticides or fungicides (Indicator IDs 233, 234), locally adapted livestock breeds (245), no use of genetically modified seeds (519) or hybrids (247) and the cultivation of some rare crop species (223).

However, there are also some low scoring subthemes. The subtheme "Greenhouse Gases" achieved one of the lowest scores of the Ethiopian systems, meaning emissions are rather high. Positively rated SMART indicators are for example no use of synthetic fertilizers or pesticides (Indicator IDs 233, 234), no use of electricity (332) and extensive pasture management (253). Negative aspects are, to name some examples of indicators with a high impact, practices prevalent in the typical systems such as the conversion of grassland to arable land (601), burning of bushes and household waste (788), and problematic practices on the arable land such as regular plowing (182), no to little mulching (237), and generally insufficient erosion prevention measures (700). Insufficient climate change mitigation actions have been pointed out by Westengen et al. (2019).

Typical Brazilian Systems

Contrary to the Ethiopian systems, the Brazilian systems show large differences in sustainability performance, as can be seen in **Figure 3**. Consequently, in this subsection, these differences

TABLE 2 | Typical coffee production systems.

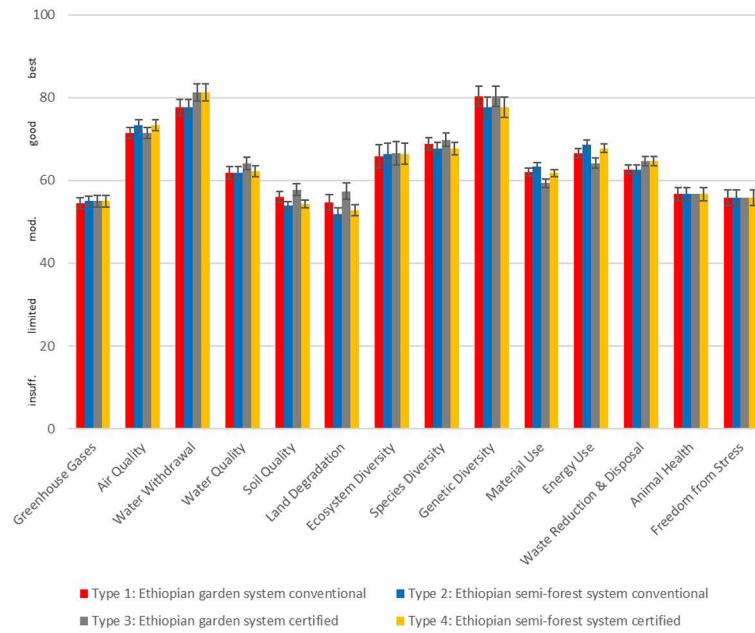| Types Classification criteria	Type 1 and 3: Ethiopian garden conventional and certified small producer	Type 2 and 4: Ethiopian semi-forest conventional and certified small producer	Type 5: Brazilian conventional large producer	Type 6: Brazilian certified small producer
VSS	• None, member of non-certified cooperative (Type 1 and 2) • Organic and Fairtrade, member of certified cooperative (Minten et al., 2015) (Type 3 and 4)		• Conventional • Most are organized in cooperatives (Carvalho, 2017)	• Organic and Fairtrade certified • Organized in certified cooperatives (Pedini, 2016)
Coffee species	• *Coffea Arabica*: • Wild growing seedlings are collected and redistributed (Gole, 2015) • Landraces, partly coffee berry disease resistant (Hylander, 2016) • Seedlings grown in own nursery, bought or received from government (Jena et al., 2012)		• *Coffea Arabica* • Propagation: Often own seedling nursery if 100 hectares or bigger (Resende, 2016; Carvalho, 2017)	• *Coffea Arabica* • Propagation: Bought-in seedlings from local markets (Pedini, 2016)
Specifications of the geographical area of interest	• Oromia and the Southern Nations, Nationalities, and People Regions (SNNPR) in the South and West of the country (Tefera and Tefera, 2014) • Precipitation between 1,500 and 2,500 mm and temperatures between 15–25°C (Tefera and Tefera, 2014) • Altitude: 1,200–2,750 meters above sea level (Tefera and Tefera, 2014)		• Minas Gerais: Sul de Minas (Vilela and Rufino, 2010) • Hilly area, very good growing conditions, good soil, enough rain (Secretaria de Produção e Agroenergia, 2009) • Altitude: 559–2,362 meters above sea level (Vilela and Rufino, 2010)	
Farm size	• 0.5 ha land (Tefera and Tefera, 2013) • 70% coffee area, 5–20 fruit trees and Ensete, 10–25% vegetable garden (Merdassa, 2016) • Primary source of income are coffee sales (Merdassa, 2016)	• 1–1.5 ha coffee forest, 0.5 arable land (Merdassa, 2016) • Plot in the forest, often not directly nearby the homestead/agricultural plots, but clear ownership structure (unlike forest coffee) (Minten et al., 2014) • Primary source of income are coffee sales (Merdassa, 2016)	• Big farms (> 50 ha), 150 ha is a typical size (Carvalho, 2017)	• Family farms (< 10 ha), 5 ha is a typical size (Pedini, 2016)
Labor	• Family labor (Hylander, 2016) • Common: Informal, reciprocal, unpaid help of neighbors in harvest time (Merdassa, 2016) • Some family child labor (Minten et al., 2015)	• Family labor (Merdassa, 2016) • 1–2 hired workers for the harvest season of 2–4 months (Merdassa, 2016) • Some family child labor (Minten et al., 2015)	• 20–30 full time employees, 10–70 seasonal workers (Carvalho, 2017)	• Family labor • One to two temporary harvest workers (Pedini, 2016)
Yield	• 400–500 kg/ha (Gole, 2015)	• 300–400 kg/ha (Gole, 2015)	• 1,800 kg/ha (Carvalho, 2017)	• 1,200 kg/ha (Moreira, 2016)
Crops and agricultural practices	• Mixed agroforestry system (Gole, 2015) • Shade trees: Fruit trees, *Ensete ventricosum*, and others (Gole, 2015) • Manual weeding 2–3 times a year (Gole, 2015) • Tree species: 5–10 (Gole, 2015) • Local cereal crops and vegetables as part of agroforestry system (Merdassa, 2016) • Low to no use of chemical pesticides and fertilizers irrespective of certification (Jena et al., 2012; Minten et al., 2015; Tefera, 2015) • Reliance completely on manual labor (Mekonnen, 2016)	• Reduction of natural forest composition to tall tree canopy of few, mostly leguminous shade trees, and coffee layer with limited number of intermediate canopy layers (Gole, 2015) • Clearing 2 times per year (Gole, 2015) • Tree species: ~19 (Gole, 2015) • Arable land: Separate cultivation of local cereals and vegetables (Merdassa, 2016)	• Plantation • Specialization in coffee (Pedini, 2016; Resende, 2016) • High level of mechanization as far as possible (Carvalho, 2017) • Weed control: Herbicides twice a year (Resende, 2016) • Inputs: Fungicides 2–3 times a year against coffee rust, as well as insecticides against coffee berry borer (Resende, 2016) • Mineral fertilizers (NPK) applied 3 times a year (Resende, 2016)	• Plantation (Pedini, 2016; Resende, 2016) • Specialization in coffee (Pedini, 2016; Resende, 2016) • Medium level of mechanization, portable harvest, and mowing machine (Moreira, 2016) • Weed control: Mechanical (Moreira, 2016) • Bought-in seedlings, rock phosphate, residues (cake) of oil production with *Ricinus communis L.*, application 2–3 times per year, together with organic material from grasses and coffee pulps (Moreira, 2016) • 30–40% use copper fungicides (Moreira, 2016)

Livestock	• Different kinds of livestock (cattle, sheep, goats, equines, and poultry) are kept in the system (Abebe, 2013)
	• Included in assessment: 70–80% extensive beef cattle free roaming on communal land (Dermauw, 2016)
	  ◦ Main use: Cultural reasons, meat for home consumption, and occasionally for sale (Dermauw, 2016)
	  ◦ Average of 0.25 ha per animal of communal grazing area (Mengistu, 2006)
	  ◦ 2–10 animals, main breeds: Zebu: e.g., Boran, Sheko, and Abigat (Mengistu, 2006)
			• None	• Only a few small animals (chicken, pigs) (Pedini, 2016), not assessed

FIGURE 2 | Sustainability performance of typical Ethiopian coffee production systems in "Environmental Integrity".
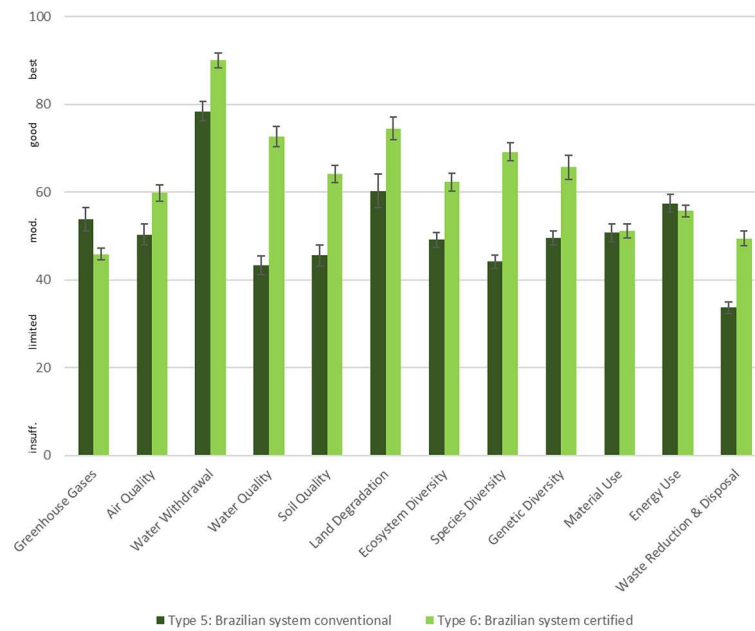

FIGURE 3 | Sustainability performance of typical Brazilian coffee production systems in "Environmental Integrity".

are mainly discussed. Agribusiness is much further developed in Brazil than in Ethiopia (Boddey et al., 2003). Here, organic certification influences the choice of inputs considerably and thus, a great sustainability improvement is visible in the environmental dimension for the typical certified system. Farm size also plays a role as the certified farms are smaller and thus not as mechanized.

The certified Brazilian farmers maintain a green cover throughout the year and only apply organic fertilizers and crop residues according to expert judgements. For pest management, they work partly with coffee rust-resistant varieties, with copper fungicides used only in rare cases. These aspects lead to a positive indicator rating and resulting good sustainability performance in the environmental dimension.
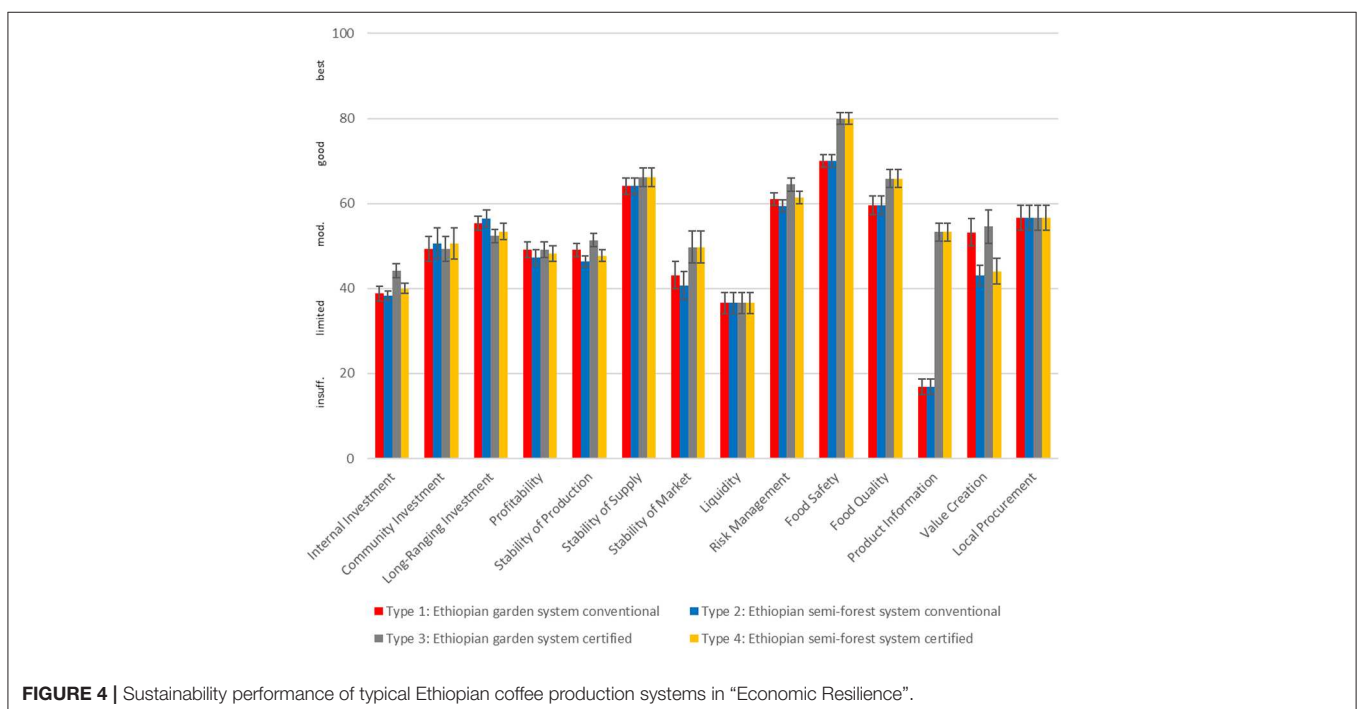
This stands in contrast to the Brazilian conventional system's scores that mostly remained in the moderate category. One of the major reasons for this discrepancy that experts mentioned is the use of synthetic pesticides and fertilizers in the conventional system. Additional negative performance is because some of the regularly used substances are considered chronically toxic, toxic to bees and aquatic organisms and persistent in soil and water by the PAN pesticide database. Experts mentioned further that heavy machinery is also likely to be used that can cause severe soil compaction.

In this paragraph, some exemplary subthemes are further explored: The conventional Brazilian system scores lowest in "Waste Reduction and Disposal" mostly due to low recycling rates in the area in general (335.1, 334, 334.1–334.5) as well as the occurrence of rather high-risk wastes related to synthetic pesticides and fertilizers (327). Boddey et al. (2003) confirm the excessive use of such substances in the conventional Brazilian coffee production, especially where the application is carried out with the help of machines. Carvalho (2006) point out the environmental risks related to the use of such substances in coffee production, such as pollution of natural resources, and the resulting negative long- and short-term health effects for humans and other living beings. Boddey et al. (2003) present organic agriculture as a solution to these problems in Brazilian coffee production. However, the certified system also has some low scoring subthemes. One of the lowest scores of the certified Brazilian system is "Greenhouse Gases." Indicator ratings with a high impact (40% and more) that cause this moderate rating are, e.g., no areas of permanent grasslands (222) or agroforestry (202) in the system and no use of fuel made from renewable sources (348) or home-produced (188) fuel. In fact, this is the only subtheme where the conventional Brazilian

system outperforms the certified system in this dimension. This is mostly due to positively rated indicators relating to the seedling nurseries used by the conventional systems, which are unlikely to use peat (733). This result is further discussed in section Conclusions.

Sustainability Performance in Dimension "Economic Resilience" on Subtheme Level
Typical Ethiopian Systems

Firstly, the overall performance of Ethiopian systems (**Figure 4**) is discussed in this subsection. Some contradicting trends are highlighted, as well as the major differences between certified and non-certified systems. Some empirical studies (Milford, 2004; Philpott et al., 2007; Dörr, 2009; Kodama, 2009) showed that certification could improve returns to smallholder coffee farmers. Several studies that have been conducted somewhat later indicate, however, that income increases through certification are generally modest (Valkila, 2009; Valkila and Nygren, 2010; Jena et al., 2012; Ruben and Fort, 2012). In many subthemes, such as "Internal Investment," "Profitability," and "Liquidity," the performance of all Ethiopian systems is not particularly high, reflecting the vulnerability to poverty of an Ethiopian smallholder highly dependent on coffee and easily affected by price changes as described, e.g., in Woubie et al. (2015) and Jena et al. (2012). This performance cannot be increased through certification in most subthemes. This finding is confirmed by several studies (Jena et al., 2012; Minten et al., 2015; Woubie et al., 2015). The main reasons in literature for these findings are, firstly, that the cooperatives can only buy a limited amount of coffee due to cash constraints. Farmers therefore sell most of their coffee to private buyers as conventional produce. Secondly, the price premium obtained by selling organic produce is only transmitted

FIGURE 4 | Sustainability performance of typical Ethiopian coffee production systems in "Economic Resilience".

to around a third to the farmers because it is captured at the cooperative level. This amount is too low to significantly improve the farmer's socio-economic situation. Thirdly, yields in cash crops in Sub-Saharan Africa remain low (Morel et al., 2019).

In this paragraph, some exemplary subthemes are further explored: For example, the driving forces (indicators with a high impact on subtheme outcome) behind the low score in the subtheme "Liquidity" are: the farmers are not able to cater for their cash needs (Indicator ID 770), no insurance against natural disasters (156) and coffee as the only income source (158). At the same time, not all subthemes of the Ethiopian systems score low. Four are rated well with "Stability of Supply" scoring highest. The good performance of this subtheme is mostly due to the system's independence of external inputs (231, 233, 234, 323, 324, 626, 199, 712). Generally, only seeds and seedlings are occasionally bought on local markets or provided by the government or cooperatives, otherwise the system is self-sufficient. Jena et al. (2012) verify these findings as the coffee farmers in their study area also mostly use very few inputs.

"Product Information" scores lowest for all conventional systems. The indicators causing this low score concern no certification of sales products (63, 65) nor of inputs (4, 5), no direct marketing (141), and generally a low transparency (175). Here, the certified systems clearly score higher. This is mostly due to the certification of the coffee and the resulting improvement of traceability. Minten et al. (2014) elaborate on traceability problems caused by the centralized trading system in Ethiopia. Certified coffee is mostly traded through cooperatives which can by-pass the Ethiopia commodity exchange (ECX) and thus traceability can be ensured (Minten et al., 2015).

Additionally, "Food Safety" and "Food Quality" show an improvement in sustainability performance in all the certified systems compared to the conventional systems, because it is more likely that measures are taken in case of a reported contamination (169) according to experts. Moreover, again improved traceability (4, 5, 63, 65) also leads to a better performance.

Lastly, substantial differences in performance can be seen concerning the subtheme "Stability of Market" in the case of the Ethiopian certified systems as they score higher than their conventional counterparts. The reasons for that are improved transparency (175) and somewhat better access to advisory services (703). This difference can however not be detected in at least 950 of 1,000 simulation runs and is not backed up with literature sources. Minten et al. (2015) do not find a difference in access to training between organic and Fairtrade certified farmers in the coffee growing regions in South-West Ethiopia. Jena et al. (2012) report only an insignificant improvement of training access for Fairtrade and/or organic certified farmers as compared to non-certified farmers in their study carried out in the Jimma region in South-West Ethiopia. In addition, they highlight the existence of a cooperative effect in Ethiopia rather than a certification effect. If a cooperative functions well, the members profit in terms of access to extension services, credit and price premiums.

Typical Brazilian Systems

As regards the Brazilian systems, some effects of certification can be seen (**Figure 5**) where the certified system scores better than the conventional. Effects of farm type can be observed in some subthemes, where the large-scale and intensified conventional Brazilian farm type scores best, presumably profiting from economies of scale. For example, "Profitability" shows the best performance for the Brazilian conventional system ranking in the category "Good." It is also the only subtheme that scores substantially higher in the conventional system than in the Brazilian certified system. The higher performance of this subtheme regarding the conventional system results mainly from the fact that, according to the expert indicator ratings, there is intensive use of synthetic fertilizers and pesticides (e.g., 233, 234) and is thus likely to obtain a high yield as opposed to the certified system. De Almeida and Zylbersztajn (2017) show that conventional farmers have a profit-oriented farming approach and are likely to obtain good prices for their produce in Brazilian large-scale coffee production. Smaller farmers are not as well organized with respect to profit optimisation. However, the certified system also shows a satisfying performance in the economic dimension. A study on the socio-economic sustainability of organic coffee farms in Brazil shows that family-run organic farms are most likely to be socio-economically sustainable. For larger organic farms, the high cost of hired labor causes severe economic constraints (Wegner et al., 2013).

Similarly to Ethiopian systems, the certified Brazilian system scores substantially higher in the subthemes "Product Information," "Food Safety," and "Food Quality." Some of the reasons mentioned by experts are the absence of harmful substances, good storage facilities, good traceability structures (4, 5, 63, 65) and few cases of contamination (169). Lastly, in contrast to the Ethiopian systems, no substantial differences in performance can be seen for the Brazilian systems concerning the subtheme "Stability of Market." For both Brazilian systems, indicator ratings show that extension services are sufficiently available (703) and thus perform as well as the Ethiopian certified systems and with less probability of variation in the outcome. However, there is evidence that coffee producers in Minas Gerais choose to be certified in order to get better access to extension services (Lemeilleur et al., 2019).

Sustainability Performance in Dimension "Social Well-Being" on Subtheme Level
Typical Ethiopian Systems

As regards the Ethiopian systems, sustainability performance generally ranges from "insufficient" to "moderate" (**Figure 6**). There are some very low scoring subthemes, which are often related to labor conditions. Labor relations are casual and there is no social security for family workers or hired workers, according to expert judgements. An example as to how this affects the performance of the subtheme "Freedom of Association and Right to Bargaining" is explained in the following. The performance of the Ethiopian garden coffee system is the lowest in this subtheme. As there is no external labor in the garden system, the indicators relevant in this subtheme only relate to working
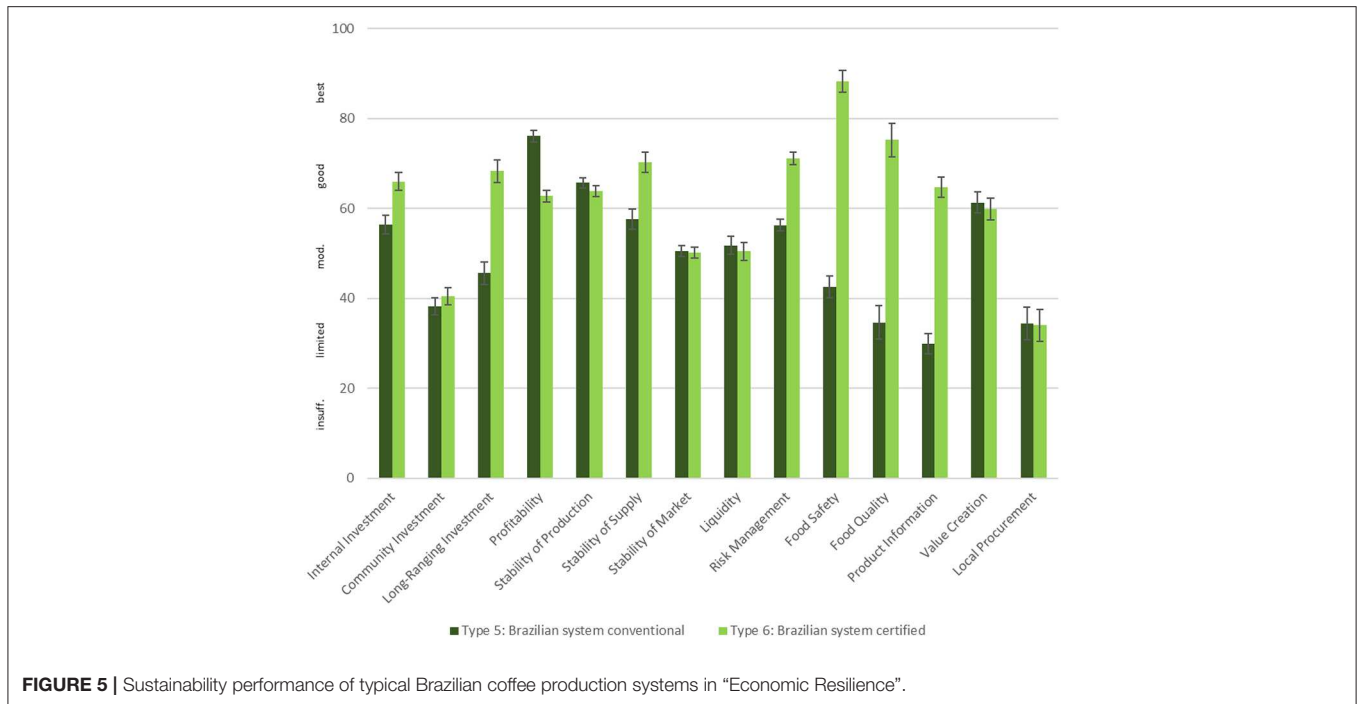
FIGURE 5 | Sustainability performance of typical Brazilian coffee production systems in "Economic Resilience".
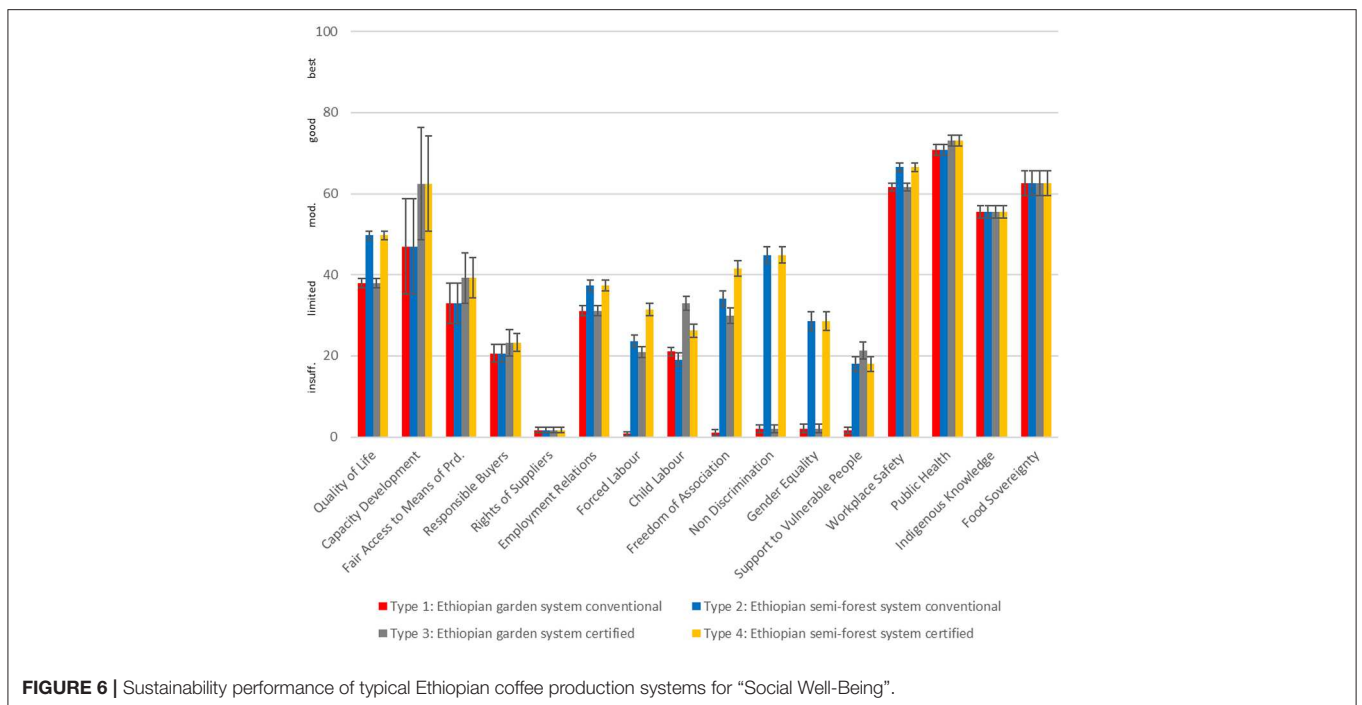

FIGURE 6 | Sustainability performance of typical Ethiopian coffee production systems for "Social Well-Being".

conditions at suppliers. The farmers do not have a socially responsible procurement strategy (5), no social certification (65) and source inputs from Ethiopia where social conditions are potentially problematic according to the International Labor Organization (ILO) (514), this subtheme scores very low. The certified garden coffee system achieves a better score due to the certification of the coffee (65). The semi-forest coffee systems

also achieve a substantially improved performance because they include positively rated indicators concerning bargaining rights of hired workers (442, 442.1).

With regard to "Non-Discrimination," "Gender Equality," "Forced Labor," and "Support to Vulnerable People," the situation is similar. In the case of "Non-Discrimination" for example, the negative indicator ratings affecting the performance of this

subtheme are the same for all systems, such as the lack of clear ownership rights (456.5) and potential socially problematic inputs (514). However, both semi-forest systems score higher, as the indicator rating includes a few more positively influencing indicators regarding external labor [e.g., no harassment of employees or forced external labor on the farm (445)]. The overall low performance of Ethiopian systems in "Gender Equality" reflects that in cash crops such as coffee, inequality between men and women in developing countries is particularly high (Tavenner et al., 2019).

Although the overall performance of the Ethiopian systems is not convincing, some subthemes score well. For example, all Ethiopian systems score highest in the subtheme "Public Health." Reasons for this good performance are, amongst others, indicator ratings indicating no or little use of synthetic pesticides (232, 233, 234) or fertilizers, antibiotics (352, 295), technologies such as GMO (519), and low amounts of waste production (e.g., 327).

Typical Brazilian Systems

Regarding the typical Brazilian systems, the certified system scores better in subthemes that relate to measures against discrimination, whereas the non-certified system shows a better performance with regard to labor relations (**Figure 7**). The good performance in the subthemes "Support to Vulnerable People," "Non-Discrimination," and "Gender Equality" of the certified Brazilian system stands in contrast to the conventional Brazilian system that performs substantially lower in these subthemes. According to expert judgments, this is due to awareness raising measures of the certified cooperatives that effectively influence farmers' attitudes toward gender equality (455, 456). According to the experts, they are likely to take measures against discrimination of women such as unequal

wages. Conventional farms do not have similar measures in place. The Coffee Barometer 2014 highlights the importance of women in coffee production making up about 50% of the work force (Panhuysen and Pierrot, 2014). Waltz (2016) emphasize that there is still a lack of empowerment of women in Southern Brazil, especially in a family farm context. The improvement in sustainability performance in the certified system in this subtheme can be confirmed by evidence from Minas Gerais in Brazil and Mesoamerica. Nelson and Pound (2009) report in a case study that a Fairtrade certified cooperative in Minas Gerais significantly improved women's participation in decision making on cooperative level through anti-discrimination measures. However, this effect could not be seen in some other cases in the area. In addition, a study assessing gender equity comparing conventional and Fairtrade/organic double certified coffee farmers in Mesoamerica shows that the certification brings significant improvements to women's control over farm practices, cash access and access to network benefits (Lyon et al., 2010). Like the Ethiopian systems, the certified Brazilian system scores highest in the subtheme "Public Health." The substantially worse performance of the conventional Brazilian system is mostly a result of indicator ratings indicating its intensive use of synthetic pesticides and fertilizers.

Regarding some labor related subthemes, the Brazilian conventional system scores either as well as or better than the Brazilian certified system. The latter is true for "Employment Relations" where the conventional system scores highest. According to expert judgments, conventional large-scale farms in the South of Minas Gerais often employ a number of workers permanently (463.1) and may pay salaries above the minimum wage to skilled laborers (410). On the contrary, in the certified Brazilian system, most work is done by the family. Occasionally,
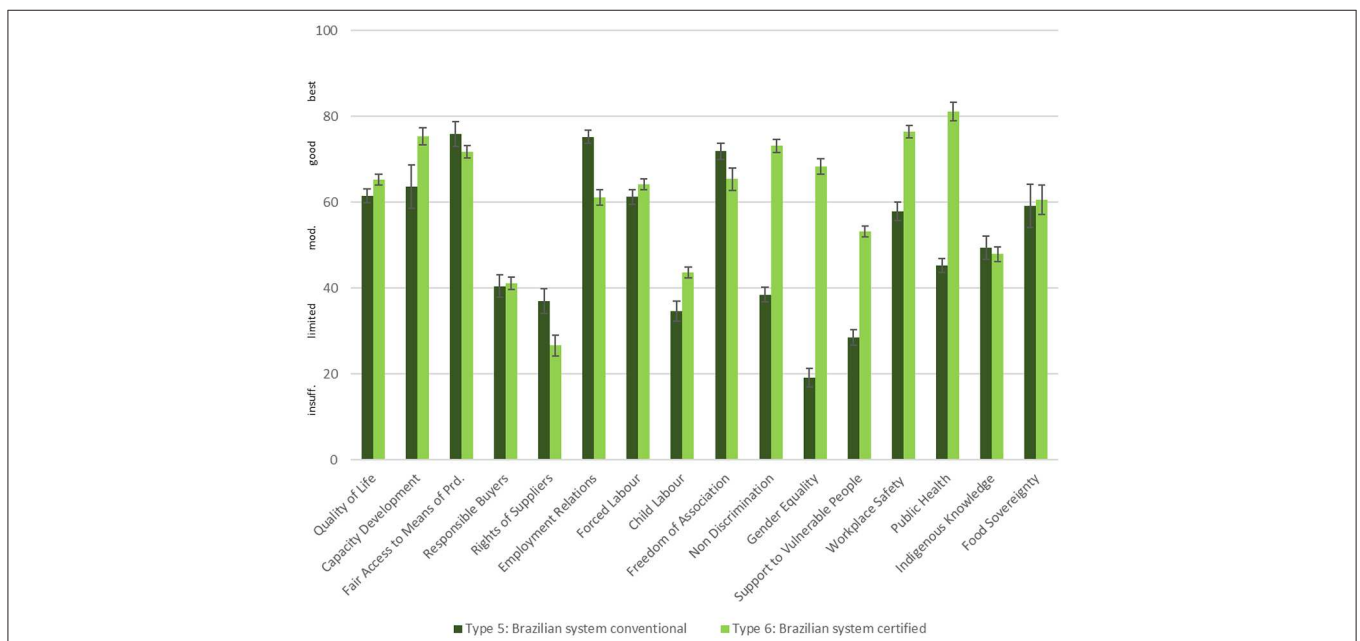

FIGURE 7 | Sustainability performance of typical Brazilian coffee production systems for "Social Well-Being".

harvest help are temporarily employed (463.1). Here, the labor conditions are also regulated, but not as well as in the big farms as the workers are not permanently employed. All employees in both systems have legally binding contracts (423), regulated working hours (437, 490), bargaining rights (442), and the right to join a union (442.1). De Almeida and Zylbersztajn (2017) support these findings in their study on success factors in the Brazilian coffee agri-chain with a focus on Minas Gerais. They point out that highly mechanized and large-scale coffee farms in Minas Gerais use skilled labor, invest in training and offer differentiated salaries. They also emphasize that such big farms are more likely to employ permanently. This stands in contrast to their characterization of Brazilian small-scale producers who are reported to rely mostly on family and temporarily hired, low skilled labor. Nevertheless, employees always have legally binding contracts and social insurance, whether permanently employed or not.

Sustainability Performance in Dimension "Good Governance" on Subtheme Level
Typical Ethiopian Systems
The performance of the Ethiopian systems mostly ranges between the categories insufficient and limited in this dimension, as can be seen in **Figure 8**. As can be seen in **Figure 5**, the subtheme "Holistic Audits" ranks the lowest for the Ethiopian non-certified systems as almost all indicators are negatively rated. According to expert indicator ratings, none of the monitoring options on the sustainability performance of the farm are used by conventional Ethiopian coffee farmers. For example, no soil samples are taken to determine fertilizer requirements (290), the farmers do not source inputs in a socially or environmentally responsible way

(4, 5) and the produce is not certified according to a social or environmental standard (63, 65). The certified Ethiopian systems perform substantially better as they are inspected internally and occasionally externally according to their certifications (63, 65).

Looking at "Remedy, Restoration & Prevention", the conventional Ethiopian systems score insufficiently as minor infringements of the law may happen (53) and no communication or conflict resolution procedures are in place (22, 28). We observe a better performance of the Ethiopian certified systems due to somewhat better supervision through certification procedures. This regards indicator ratings showing that restriction measures against infringements of the law (53) such as extending the agricultural area into a nearby forest and measures against contamination of produce (169) are in place. The Ethiopian semi-forest systems score the lowest of all systems in the subtheme "Legitimacy" as the employment situation in these systems is precarious (e.g., 423, 410) according to expert judgements.

The Ethiopian systems all score the same in the subtheme "Mission Statement," in the category "Limited." Typical Ethiopian coffee farmers are partly committed to sustainability topics (8), but cannot name specific planned improvements (750). No difference can be found between the certified and the non-certified systems as the farmers are often not aware of their cooperative's certification and its meaning (Jena et al., 2012).

Typical Brazilian Systems
There are several other subthemes where the certified system outperforms the non-certified system. "Transparency" e.g., scores rather low for the conventional system (**Figure 9**). Through certification and the resulting requirements on a better traceability (165, 63, 65), the rating of the certified
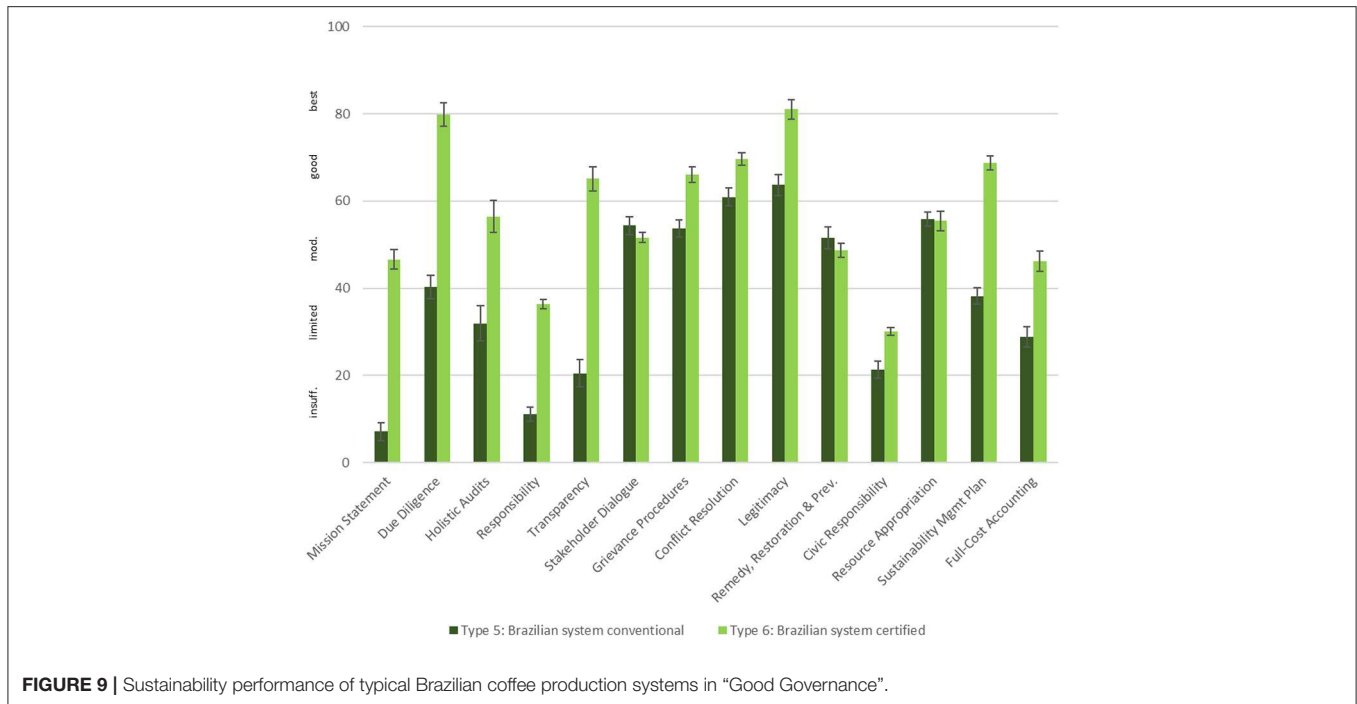

FIGURE 8 | Sustainability performance of typical Ethiopian coffee production systems in "Good Governance".
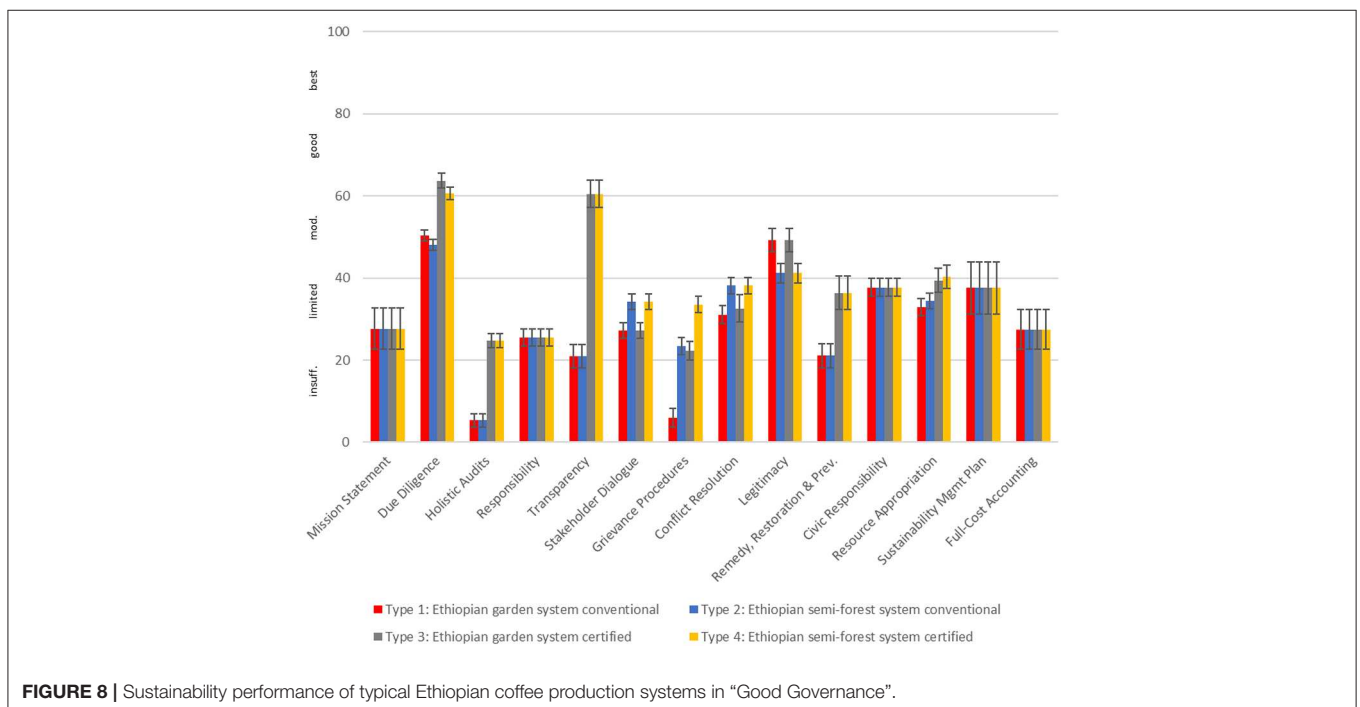
FIGURE 9 | Sustainability performance of typical Brazilian coffee production systems in "Good Governance".

system is substantially better. Another example is the subtheme "Mission Statement," where the conventional Brazilian system is outperformed by the certified system. This is partly due to the different sizes of the farm types. The SMART -Farm Tool omits some indicators regarding written commitments to sustainability and the publication of such material for a smallholder. According to expert indicator ratings, typical certified smallholders do commit themselves verbally to sustainability issues (8). In the SMART-Farm Tool it is assumed that the typical conventional farm as a bigger enterprise has the resources to issue something like a sustainability report (6, 35). However, the typical conventional farm is not likely to have any such documentation. This explains its very low performance in "Mission Statement."

The Brazilian systems score similarly and substantially better in the subtheme "Remedy, Restoration & Prevention" than the Ethiopian systems. Both are unlikely to be involved in infringements (53) of the law as regards labor regulations or conflicts with neighbors (22) according to expert indicator ratings. Expert indicator ratings also suggest that they take measures in cases of product contamination (169). This also explains to a good extent the performance of the Brazilian systems in the subtheme "Legitimacy" where both systems obtain the highest degree of goal achievement of all subthemes in this dimension (Category "Good" in the conventional system and even "Best" in the certified system). It shows that, according to expert judgements, both systems are mostly compliant with the applicable national laws and international human rights standards. The certified system scores better because it is unlikely to cause any negative social or environmental impacts whereas the conventional system my cause some environmental pollution due to intensive practices (21), such as pesticide use and resulting environmental pollution.

Discussion of the Method and Study Limitations

Typical coffee production systems at the country level in Ethiopia and Brazil could be distinguished and successfully assessed with the SMART-Farm Tool. One of the main criteria to select the typical farm types was the production and export amount of coffee. In Ethiopia, certification status and farm size can be well isolated, thus certification effects could be estimated. However, in Brazil farm size and certification are correlated, therefore we had to choose a certified smallholder system and a conventional large-scale farm. The results should be viewed in context of this, and any certification effect treated with caution. An advantage of this method is however that a realistic picture of an existing farm type is drawn instead of attempting to create an artificial counterfactual that does not reflect the typical situation.

The method of collecting data through expert interviews is generally considered an easy and efficient way to obtain information (Mieg and Näf, 2005; Glaser et al., 2010; Bogner et al., 2014).

However, in this study some difficulties arose when discussing the SMART indicators for the typical systems we defined. As the questionnaire is originally designed for a real farm assessment, experts occasionally found it difficult to give an estimation for a typical farm. In general, extracting data in such detail as required by the SMART-Farm Tool was challenging. Some examples are indicators such as "Settings of combustion motors [How often are the settings of combustion motors of vehicles (e.g., tractor, stapler) and other machineries checked and adjusted (engine, air filter etc.)?]" or "Infringements of the law (In the last 5 years, have there been any cases in which the farm has broken the law? If yes, how serious were they?)." Such questions could

be omitted or replaced by more general indicators for future expert judgements.

On the contrary, single farm assessments often are prone to biased answers by farmers, particularly if sensitive issues such as pesticide use, child labor, forced labor, infringement of the law, and gender equality are concerned. Independent experts may have a more objective view. An example is a recently published study by Ssebunya et al. (2019) where no child labor was found in coffee production systems during sustainability assessments of individual farms using SMART. However, it is likely to be prevalent according to Akoyi et al. (2018). Following the approach of interviewing experts can thus yield complementary information for sustainability assessments in which numerous sensitive topics are analyzed. Issues such as social desirability and conformity bias, as well as prestige bias can be circumvented.

Apart from the content side, we see a methodological contribution of this paper as there are so far only very few studies, which follow such a structured expert-based approach, and then validate it with accounts from the literature. Furthermore, the concept of looking at typical farms that represent the mode of a distribution rather than the mean, has several advantages. Firstly, the combination of farm characteristics of the typical farms can be observed in the field and thus specific recommendations can be made. This stands in contrast to studying an "artificial" average farm where the modeled combination of farm characteristics cannot be found in the field. Here, an aggregation bias is likely where we assume that the average farm has a larger option space than farms in the field really have (Feuz and Skold, 1992; Häring, 2003a). Secondly, economizing of resources used in the study while still yielding relevant results can be consolidated in this approach (Feuz and Skold, 1992). This is especially relevant for sustainability assessments, where a large amount of data in the different dimensions needs to be collected per farm.

A promising approach and thus a topic for further research could also be a combination of farm assessments and expert interviews to apply SMART at the sectoral level. Here, the advantages of both approaches could be capitalized on and a robustness check of results could be conducted. In order to efficiently coordinate this process, some farms fulfilling the description of the defined typical farm could be assessed individually.

During the process of data collection, the question of bias caused by the manner of interviewing or the background of the experts arose several times. Bogner et al. (2014) confirm that complete neutrality in expert interviews is methodologically impossible. They argue that there are, e.g., different ways the interviewer can be perceived by the interviewee, causing him, or her to give differing answers. The methods used in this study aiming to avoid a potential bias are: (1) the number of interviewed persons from groups with strong political or commercial interests were kept to a minimum, (2) the interview procedure as described in the Methods section was followed as an attempt at standardization, (3) uncertainties in indicator ratings were accounted for through Monte-Carlo simulations. Nevertheless, neutrality of the answers cannot be entirely ensured due to the nature of the data.

As mentioned in Feuz and Skold (1992), the selected farm types are not representative. In addition, the data collection for indicator rating was conducted in a qualitative way. This means that the results of this study cannot be claimed to have any statistical representativeness, but rather give an overview of perceived typical situations in the field. This needs to be kept in mind when interpreting the results.

The choice of an appropriate sustainability assessment tool is crucial. Nowadays a wide variety of sustainability assessments and tools are available (FAO, 2014). The SMART-Farm Tool is based on a multi-criteria assessment approach designed to assess the sustainability performance of a farm with relatively low cost and based on the data easily available at farm level. Schader et al. (2016) mention that for some subthemes such as "Energy Use," "Greenhouse Gases," and "Profitability," a more quantitative method like a life cycle assessment or the calculation of gross margins may be alternative assessment methods. However, they argue that these approaches are costlier as the data may either not be available or the farmers may be hesitant to disclose them, especially in the case of economic data. Nevertheless, they also point out that a further in-depth adaptation of the pool of indicators in the SMART-Farm Tool from case to case may be advisable. During this research, several areas were identified where improvements would be helpful for future assessments of similar production systems. Some examples are indicators addressing agroforestry and perennial crop production characteristics more in depth as well as price spreads and volatility. Additionally, an extension of the tool from farm level to organizational or processing level may be of use in the coffee production context.

Finally, there is some distortion of results caused by the number of indicators relevant for each system. For example, the conventional Brazilian system scores better in the subtheme "Greenhouse Gases" than the certified system. This is caused by the fact that the former contains indicators accounting for practices in seedling nurseries. However, it is likely that practices are similar also in the external seedling nurseries from which the certified system buys. This increase in sustainability is thus not real, but rather a construct of the method setting the system boundaries as such. A similar case can be observed when evaluating the Ethiopian systems with and without hired labor. e.g., in the subtheme "Freedom of Association and Right to Bargaining," both systems score the same except from two additional indicators only relevant for the system using hired labor. This leads to a situation where the two systems are not easily comparable. This is a model-inherent issue that will need to be addressed in the future, for example by selecting generic indicators relevant for each system for comparison.

Overall, it remains a challenge to find a balance between a more individualized approach that reflects the specific characteristics of the described system well, and a method that still grants comparability across many different systems. Here, trade-offs depending on the intention of use of the respective sustainability assessment in a specific context cannot be avoided entirely.

Lastly, Monte-Carlo simulations were already used successfully in a similar context in order to calculate uncertainty

distributions of the SAFA subthemes if the indicator weights are uncertain (Schader et al., 2019). In this study, the method proved itself helpful to take variations within the typical systems defined into account as well as the uncertainty resulting from the respective data source. With regard to the additional uncertainty, Weidema et al. (2013) mention that the uncertainties they estimated may have been understated. Hence, variation within the subthemes may be even larger than depicted in this study.

CONCLUSIONS

This study evaluated the sustainability performance of typical certified and non-certified coffee production systems at indicator level. In the following, we answer the two research questions asked in the introductory section for each of the four dimensions of sustainability. Firstly, the differences and similarities in performance between the typical certified and conventional systems are highlighted. Secondly, differences and similarities between the country cases are shown for each dimension of sustainability.

In the environmental dimension, organic, and Fairtrade certification do not show an impact for typical coffee production systems in Ethiopia as all systems are extensively managed with low external input use. The effect of organic certification may become more visible and valued in the environmental dimension when intensification progresses in Ethiopia. The situation is different in Brazil where agribusiness is much further developed than in Ethiopia. Here, organic certification influences the choice of inputs considerably and thus, a great sustainability improvement is visible in the environmental dimension for the typical certified system. Farm size also plays a role as the certified farms are smaller and thus not as mechanized.

Regarding the economic dimension, effects of certification can be seen in the subthemes "Product Information," "Food Safety" and "Food Quality" as the certified systems score better than conventional counterparts in both countries. Effects of farm type can be observed in the subtheme "Profitability" where the large-scale and intensified conventional Brazilian farm type scores best, profiting from economies of scale. Effects of geographical and political conditions in a country are observable for the Ethiopian systems in subthemes like "Profitability" and "Liquidity" as the majority of the rural population of Ethiopia lives on incomes below the poverty line and certification is not able to lift this.

Certification positively affects the subthemes "Gender Equality," "Support to Vulnerable People," "Non-Discrimination," and "Public Health" in the social dimension regarding the sustainability performance of the Brazilian certified system. Measures for more gender equity are the main driving forces for this performance improvement for the first three

mentioned subthemes, organic practices for the fourth. A farm type effect can again be seen for the conventional Brazilian system for "Employment Relations" as the large farm is more likely to employ workers permanently. For the Ethiopian systems, this effect shows in the subthemes "Non-Discrimination," "Gender Equality," and "Freedom of Association and Right to Bargaining" where the use or non-use of hired labor mostly causes the differences in sustainability performance as labor conditions are precarious.

In the governance dimension, effects of certification can be seen for "Holistic Audits" and "Transparency" for all systems. For the Ethiopian systems, this effect is also visible for "Remedy, Restoration & Prevention," whereas the Brazilian systems score similarly as both are unlikely to be involved in infringements of the law as regards labor regulations or conflicts with neighbors. Effects of geographical and political conditions show in "Legitimacy" as the Brazilian systems are more compliant with the applicable national laws and international human rights standards than the Ethiopian systems.

The SMART-Farm Tool in combination with typical farm theory and data collection through expert interviews can give an interesting first impression on the general dynamics of sustainability in typical agricultural production systems, although it is crucial to state that the data are not statistically representative. It can thereby identify hotspots that may be addressed as well as good practices that may be implemented elsewhere by decision makers.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

AUTHOR CONTRIBUTIONS

EW was the main data collector and writer. SM has contributed substantially to the conceptualization and the technical implementation of the Monte-Carlo analyses. LB has contributed substantially to the conceptualization and the technical implementation of the sustainability assessment. MC and MS have contributed substantially to the conceptualization and the structure of the paper. CS has initiated and conceptualized the project and has been involved in every research step.

SUPPLEMENTARY MATERIAL

REFERENCES

Abdissa, F. M., de Mey, Y., Nyssen, J., and Maertens, M. (2017). Do private sustainability standards contribute to poverty alleviation? A comparison of different coffee certification schemes in Ethiopia. *Sustainability* 9, 246–267. doi: 10.3390/su9020246

Abebe, T. (2013). Determinants of crop diversity and composition in Enset-coffee agroforestry homegardens of Southern Ethiopia. *J. Agricul. Rural Dev. Trop. Subtrop.* 114, 1283–1293. Available online at: http://www.nbn-resolving.de/urn:nbn:de:hebis:34-2013030542580

Akoyi, K., Mitiku, F., and Maertens, M. (2018). *Is Prohibiting Child Labour Enough?* Coffee certification and child schooling in Ethiopia and Uganda.

Barbosa, J. N., Borem, F. M., Cirillo, M. A., Malta, M. R., Alvarenga, A. A., and Alves, H. M. R. (2012). Coffee quality and its interactions with environmental factors in minas gerais, Brazil. *J. Agricult. Sci.* 4, 181–190. doi: 10.5539/jas.v4n5p181

Blackman, A., and Naranjo, M. (2010). Does eco-certification have environmental benefits? *Organic Coffee in Costa Rica Resources for the Future Discussion Paper* 10. doi: 10.2139/ssrn.1713888

Boddey, R. M., Xavier, D. F., Alves, B. J., and Urquiaga, S. (2003). Brazilian agriculture: the transition to sustainability. *J. Crop Product.* 9, 593–621. doi: 10.1300/J144v09n01_10

Bogner, A., Littig, B., and Menz, W. (2014). *Interviews mit Experten: eine praxisorientierte Einführung*. Berlin: Springer-Verlag. doi: 10.1007/978-3-531-19416-5

Carvalho, D. G. R. (2017). *Agronomist and Advisor, Personal Correspondence on 27.07.2017.*

Carvalho, F. P. (2006). Agriculture, pesticides, food security and food safety. *Environ. Sci. Policy* 9, 685–692. doi: 10.1016/j.envsci.2006.08.002

De Almeida, L., and Zylbersztajn, D. (2017). Key success factors in the brazilian coffee agrichain: Present and future challenges. *Proc. Food Syst. Dyn.* 8, 45–53.

Dermauw, D. V. (2016). Tropical livestock scientist, Personal correspondence on 28.10.2016.

Dörr, A. C. (2009). *Economic Analysis of Certification in the Brazilian Fruit Chain.* Göttingen: Cuvillier Publisher.

European Commission (2007). *Council Regulation (EC) No 834/2007 of 28 June 2007 on organic production and labelling of organic products and repealing Regulation (EEC) No 2092/91.*

FAO (2014). *Guidelines for Sustainability Assessment in Food and Agriculture (SAFA).* Rome, Food and Agriculture Organization of the United Nations (FAO).

FAO (2016). *Sustainability Pathways.* Available online at: http://www.fao.org/nr/sustainability/sustainability-assessments-safa/en/ (accessed April 16, 2016).

FAOSTAT (2016). *Food and Agriculture Data.* Available online at: http://faostat3.fao.org (accessed April 20, 2016).

Feuz, D. M., and Skold, M. D. (1992). Typical farm theory in agricultural research. *J. Sustain. Agricult.* 2, 43–58. doi: 10.1300/J064v02n02_05

FLO (2017). *Aims of Fairtrade Standards.* Available online at: https://www.fairtrade.net/standards/aims-of-fairtrade-standards.html (accessed January 31, 2017).

Giuliani, E., Ciravegna, L., Vezzulli, A., and Kilian, B. (2017). Decoupling standards from practice: the impact of in-house certifications on coffee farms' environmental and social conduct. *World Dev.* 96, 294–314. doi: 10.1016/j.worlddev.2017.03.013

Glaser, B. G., Strauss, A. L., Paul, A. T., and Kaufmann, S. (2010). *Grounded theory: Strategien qualitativer Forschung.* Huber.

Gole, T. W. (2015). *Coffee: Ethiopia's Gift to the World.* Environment and Coffee Forest Forum (ECFF), Addis Abeba, Ethiopia.

Hardt, E., Borgomeo, E., dos Santos, R. F., Pinto, L. F. G., Metzger, J. P., Sparovek, G. (2015). Does certification improve biodiversity conservation in Brazilian coffee farms? *For. Ecol. Manage.* 357, 181–194. doi: 10.1016/j.foreco.2015.08.021

Häring, A. M. (2003a). *An Interactive Approach to Policy Impact Assessment for Organic Farms in Europe.* University of Hohenheim, Germany.

Häring, A. M. (2003b). *An Interactive Approach to Policy Impact Assessment for Organic Farms in Europe.* University of Hohenheim, Germany.

Hylander, D. K. (2016). *Natural Scientist, Personal Correspondence on 28.09.2016.*

Jena, P. R., Chichaibelu, B. B., Stellmacher, T., and Grote, U. (2012). The impact of coffee certification on small-scale producers' livelihoods: a case study from the Jimma Zone, Ethiopia. *Agricult. Econ.* 43, 429–440. doi: 10.1111/j.1574-0862.2012.00594.x

Jurjonas, M., Crossman, K., Solomon, J., and Baez, W. L. (2016). Potential links between certified organic coffee and deforestation in a protected area in Chiapas, Mexico. *World Dev.* 78, 13–21. doi: 10.1016/j.worlddev.2015.10.030

Kodama, Y. (2009). "The effects of Fair Trade on coffee producers: a case study of Ethiopian coffee cooperatives," in *16th International Conference of Ethiopian Studies, Trondheim,* eds S Ege, HAB Teferra, and B Shiferaw (Norway), 297–299.

Lemeilleur, S., Julie, S., Presoto, A. E., Souza, R. D. C., and Saes, M. S. M. (2019). *Coffee Farmers' Motivations to Comply With Sustainability Standards.* CEE-M, Universtiy of Montpellier, CNRS, INRA, Montpellier SupAgro.

Lernoud, J., Potts, J., Sampson, G., Vivek, V., Willer, H., and Wozniak, J. (2016). *The State of Sustainable Markets - Statistics and Emerging Trends.* International Trade Centre (ITC).

Lyon, S., Bezaury, J. A., and Mutersbaugh, T. (2010). Gender equity in fairtrade-organic coffee producer organizations: Cases from Mesoamerica. *Geoforum* 41, 93–103.

Mekonnen, M. (2016). *Natural Scientist, Personal Correspondence on 27.09.2016.*

Mengistu, A. (2006). *Country Pasture/Forage Resource Profiles.* Ethiopia, FAO.

Merdassa, J. (2016). *Economist, Personal Correspondence on 24.10.2016.*

Mieg, H. A., and Näf, M. (2005). *Experteninterviews in den Umwelt- und Planungswissenschaften.* Eine Einführung und Anleitung. ETH Zürich.

Milford, A. (2004). *Coffee, Co-Operatives and Competition: the Impact of Fair Trade.* Bergen: Chr. Michelsen Institute.

Minten, B., Dereje, M., Engeda, E., and Tamru, S. (2015). *Who Benefits From the Rapidly Increasing Voluntary Sustainability Standards?* Evidence from Fairtrade and Organic certified coffee in Ethiopia. Ethiopia Strategy Support Program. Ethiopian Development Research Institute (EDRI), International Food Policy Research Institute (IFPRI), 1–25.

Minten, B., Tamru, S., Kuma, T., and Nyarko, Y. (2014). *Structure and Performance of Ethiopia's Coffee Export Sector.* Ethiopia Strategy Support Program. Ethiopian Development Research Institute (EDRI), International Food Policy Research Institute (IFPRI), 1–30.

Moreira, D. C. F. (2016). *Agronomist.* Personal Correspondence. (accessed December 13, 2016).

Morel, A. C., Hirons, M., Adu Sasu, M., Quaye, M., Ashley Asare, R., Mason, J., et al. (2019). The ecological limits of poverty alleviation in an african forest-agriculture landscape. *Front. Sustain. Food Syst.* 3:57. doi: 10.3389/fsufs.2019.00057

Nelson, V., and Pound, B. (2009). The last ten years: A comprehensive review of the literature on the impact of Fairtrade. *Nat. Resour. Inst.* 1–48.

Palisade (2016). *@RISK-Software.* Available online at: https://www.palisade.com/risk/ (accessed October 10, 2016).

Panhuysen, S., and Pierrot, J. (2014). *Coffee Barometer 2014.* Coffee Barometer. Hivos IUCN Nederland Oxfam Novib Solidaridad WWF.

Pedini, P. S. (2016). *Agronomist, Personal Correspondence on 29.11.2016, 01.12.2016.*

Philpott, S., Bichier, P., Rice, R., and Greenberg, R. (2007). Field-testing ecological and economic benefits of coffee certification programs. *Conserv. Biol.* 21, 975–985. doi: 10.1111/j.1523-1739.2007.00728.x

Pierrot, J., Giovannucci, D., and Kasterine, A. (2011). *Trends in the Trade of Certified Coffees.* International Trade Centre.

Piras, E. (2011). *Reis ist Leben - Wie Öko-Landbau und Fairer Handel zu Ernährungssouveränität führen.* Naturland e. V.

Reidsma, P., Spiegel, A., Paas, W., Accatino, F., Antonioli, F., Appel, F., et al. (2019). *D5.3 Resilience Assessment of Current Farming Systems Across.* European Union.

Resende, P. M. L. V. (2016). *Agronomist, Personal Correspondence on 25.11.2016, 16.12.2016.*

Ruben, R., and Fort, R. (2012). The impact of fair trade certification for coffee farmers in Peru. *World Dev.* 40, 570–582. doi: 10.1016/j.worlddev.2011.07.030

Rubinstein, R. Y., and Kroese, D. P. (2011). *Simulation and the Monte Carlo Method.* Hoboken, NJ: Wiley. Com.

Schader, C., Baumgart, L., Landert, J., Muller, A., Ssebunya, B., Blockeel, J., et al. (2016). Using the Sustainability Monitoring and Assessment Routine (SMART) for the systematic analysis of trade-offs and synergies between sustainability dimensions and themes at farm level. *Sustainability* 8:274. doi: 10.3390/su8030274

Schader, C., Michael, C., Anja, H., Jan, L., Johan, B., Lukas, B., et al. (2019). Accounting for uncertainty in multi-criteria sustainability assessments at the farm level: Improving the robustness of the SMART-Farm Tool. *Ecol. Indicat.* 106:105503. doi: 10.1016/j.ecolind.2019.105503

Secretaria de Produção e Agroenergia (2009). *Análise Estrutural da Cafeicultura Brasileira.* Ministério da Agricultura, Pecuária e Abastecimento.

Ssebunya, B. R., Schader, C., Baumgart, L., Landert, J., Altenbuchner, C., Schmid, E., et al. (2019). Sustainability performance of certified and non-certified smallholder coffee farms in Uganda. *Ecol. Econ.* 156, 35–47. doi: 10.1016/j.ecolecon.2018.09.004

Tavenner, K., van Wijk, M., Fraval, S., Hammond, J., Baltenweck, I., Teufel, N., et al. (2019). Intensifying inequality? gendered trends in commercializing and diversifying smallholder farming systems in East Africa. *Front. Sustain. Food Syst.* 3:10. doi: 10.3389/fsufs.2019.00010

Tefera, A. (2015). *Coffee Annual Report Ethiopia.* USDA Foreign Agricultural Service.

Tefera, A. (2016). *Coffee Annual Report Ethiopia.* USDA Foreign Agricultural Service.

Tefera, A., and Tefera, T. (2013). *Coffee Annual Report Ethiopia.* USDA Foreign Agricultural Service.

Tefera, A., and Tefera, T. (2014). *Coffee Annual Report Ethiopia.* USDA Foreign Agricultural Service.

Thünen Institute (2016). *A Standard Operating Procedure to Define Typical Organic Farms.* Thünen Institute of Farm Economics, Braunschweig.

Valkila, J. (2009). Fair Trade organic coffee production in Nicaragua – Sustainable development or a poverty trap? *Ecol. Econom.* 68. doi: 10.1016/j.ecolecon.2009.07.002

Valkila, J., and Nygren, A. (2010). Impacts of Fair Trade certification on coffee farmers, cooperatives, and laborers in Nicaragua. *Agricult. Human Values* 27, 321–333. doi: 10.1007/s10460-009-9208-7

Vanderhaegen, K., Akoyi, K. T., Dekoninck, W., Jocqué, R., Muys, B., Verbist, B., et al. (2018). Do private coffee standards 'walk the talk' in improving socio-economic and environmental sustainability? *Global Environ. Change* 51, 1–9. doi: 10.1016/j.gloenvcha.2018.04.014

Vilela, P. S., and Rufino, J. L. (2010). Caracterização da Cafeicultura de Montanha de Minas Gerais. Estudos INAES. Cadeias Produtivas. Café. *Instituto Antonio Ernesto de Salvo,* 300.

Waltz, A. (2016). The women who feed us: gender empowerment (or lack thereof) in rural Southern Brazil. *J. Rural Stud.* 47, 31–40. doi: 10.1016/j.jrurstud.2016.07.009

Wegner, R., Turco, P. H. N., and Bliska, F. M. (2013). Organic agriculture: socioeconomic sustainability of Brazilian Coffee. *J. Agricult. Sci. Tech. A* 3.

Weidema, B. P., Bauer, C., Hischier, R., Mutel, C., Nemecek, T., Reinhard, J., et al. (2013). *Overview and Methodology - Data Quality Guideline for the Ecoinvent Database Version 3.* Swiss Centre for Life Cycle Inventories, St. Gallen, Switzerland.

Westengen, O. T., Haug, R., Guthiga, P., and Macharia, E. (2019). Governing seeds in East Africa in the face of climate change: assessing political and social outcomes. *Front. Sustain. Food Syst.* 3, 1–14. doi: 10.3389/fsufs.2019.00053

Woubie, A. A., Muradian, R., and Ruben, R. (2015). "Impact of multiple certification on smallholder coffee farmers' livelihoods: evidence from southern Ethiopia," in *Coffee certification in East Africa,* eds R. Ruben and P. Hoebink (Wageningen: Wageningen Academic Publishers), 127. doi: 10.3920/978-90-8686-805-6_4