UNITED NATIONS CONFERENCE ON TRADE AND DEVELOPMENT

July 2020

UNCTAD Research Paper No. 50
UNCTAD/SER.RP/2020/9

The Trade Impact of Voluntary Sustainability Standards:
A review of empirical evidence

Niematallah E.A. Elamin

and

Santiago Fernandez de Cordoba

Trade Analysis Branch
Division on International Trade and Commodities
UNCTAD

santiago.fernandezdecordoba@unctad.org

niematallah.ahmedelamin@unctad.org

Abstract

The expansion of Voluntary Sustainability Standards (VSS) has introduced tools to address key sustainability challenges and expand trade as well as new complexity in trade policy dialogues. Although VSS are voluntary, they have become a market reality and non-compliance can lead to the exclusion of producers from Global Value Chains (GVCs). Although the literature has adequately addressed the theoretical trade impact of VSS, there is a worrying lack of empirical analysis in this field.

The current research is to our knowledge the first to survey the empirical literature on the trade impact of VSS. Surprisingly we only found a limited number of empirical studies (nine studies) in this field. On the one hand, the reviewed studies give contradictory results, but in addition, are few in number and case-specific that focus on very few standards and products.

This paper aims to draw the attention of researchers towards the lack of evidence in this area. Given that the gap in the literature can partly be explained by data access, we also present some of the publicly available data sources and highlights the significance of increasing transparency in terms of data availability.

Key words: Voluntary Sustainability Standards, International Trade, Trade models, Sustainable Development

Contents

Acknowledgements

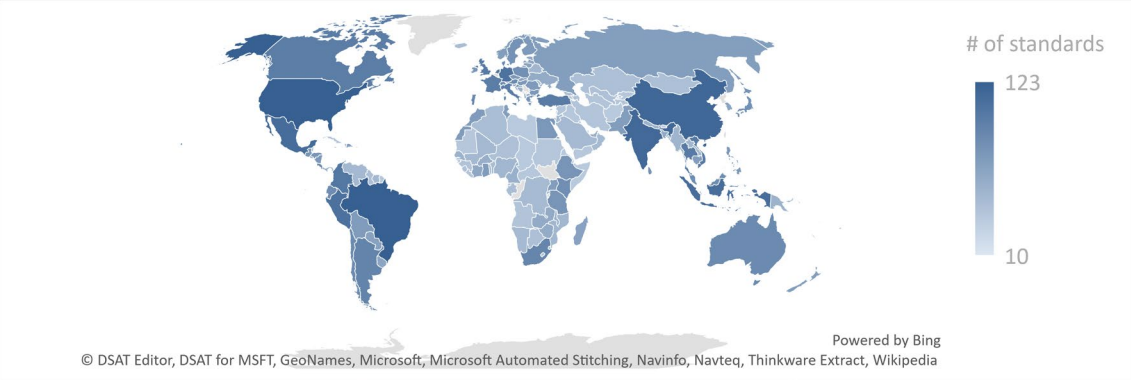
The authors are grateful to the valuable comments received from Cecilia Heuser (UNCTAD), Axel Marx (University of Leuven), and Janne Bemelmans (University of Leuven). The authors also benefit from discussions with Clara Brandi (German Development Institute (DIE)), Matteo Fiorini (European University Institute (EUI)), Katharina Bissinger (German Development Institute (DIE)), Helga Willer (The Research Institute of Organic Agriculture (FiBL)), and Richter Toralfon (The Research Institute of Organic Agriculture (FiBL)) on the subject. The study was formatted by Jenifer Tacardon-Mercado (UNCTAD).

1. Background

Over the last decades, Voluntary Sustainability Standards (VSS) have emerged as new tools to address key sustainability challenges such as biodiversity, climate change, and human rights. The United Nations Forum on Sustainability Standards (UNFSS), describes VSS as ''specifying requirements that producers, traders, manufacturers, retailers or service providers may be asked to meet, relating to a wide range of sustainability metrics, including respect for basic human rights, worker health and safety, the environmental impacts of production, community relations, land use planning and others'' (UNFSS,2013 page 3).[1] VSS schemes can be conceptualized as non-state, market-driven governance approaches that aim to improve the economic, environmental, and social sustainability of production through the establishment and enforcement of specific norms of behavior (Cashore et al., 2004). These voluntary sustainability standards are a key governance instrument to facilitate the transition to a green economy (Marx & Wouters, 2015).

The number of VSS has continued to grow; currently, there are around 264 active VSS (according to ITC Standards Map) in 194 countries and 15 sectors, and about 457 ecolabels (according to Ecolabel Index) in 199 countries, and 25 industry sectors.[2] The growth in the number of certified products is driven by the strong demand for products that are certified according to sustainability standards and is primarily driven by large retailers (UNFSS, 2013; Marx & Wouters, 2015). Also, certification assures market access specially to developed countries' markets as non-compliance can lead to the exclusion of producers from Global Value Chains (GVCs). Also, more responsible and sustainable production practices have been encouraged by the public sector, particularly in developed countries (ITC, 2017). One possible explanation of the growing demand for certified products is the increasing awareness of environmental protection and social issues.

Figure 1. Number of VSS by country



Although the number of VSS, along with the number of certified products and certified producers, are growing, this growth is not distributed evenly across the globe. Figure 1 presents the active number of VSS by country and indicates the uneven distribution of VSS between developed, developing, and least developed countries.[3] ITC (2017) states that around 74% of the standards in the Standards Map database are active in OECD countries. Countries in the Middle East and North Africa, Sub-Saharan Africa, and Central Asia have access to a lower number of standards than the global average of 33 per country (ITC, 2017). From a demand perspective,

[1] Although VSS and private standards are often used interchangeably, a distinction needs to be made. VSS include all the nongovernmental systems that are not mandatory, i.e. the compliance with them is not mandatory under international trade rules. Private standards refer to the standards set by the private sector and requested by it (not by public authorities).

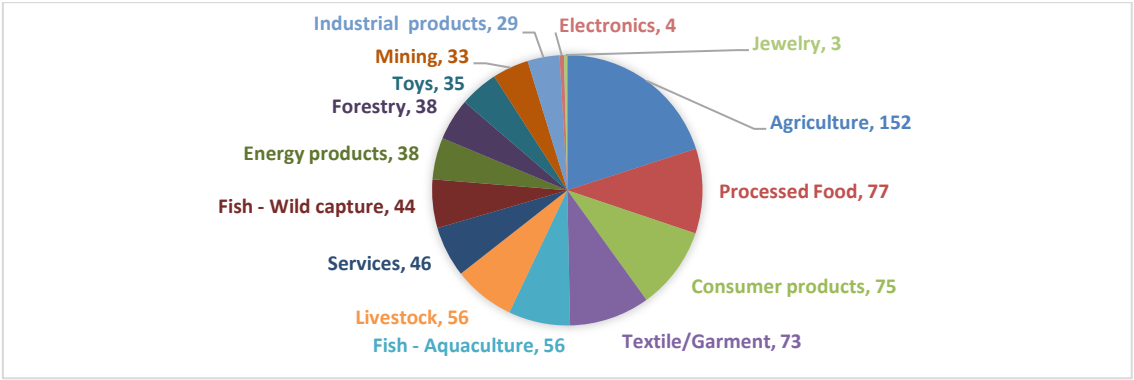
[2] ITC Standards Map and Ecolabel Index follow different methodologies in constructing their databases. The ITC Standards Map is more restrictive in recognizing and reviewing VSS. Also, the Ecolabel Index includes a significant number of corporate codes of conducts which are (by definition) not VSS.

[3] Generated using International Trade Centre (ITC). 2020. ITC Standards Map. Retrieved from: https://www.sustainabilitymap.org/standards?q=eyJzZWxlY3RlZENsaWVudCI6Ik5PIEFGRklMSUFUSU9OIn0%3D

developed countries are the main markets for certified products with more than 95 per cent of sales (Liu, 2009). FiBL (2018) states that 90 per cent of the organic food and drinks sales are concentrated in North America and Europe.

VSS are used in global supply chains in various sectors and industries such as agriculture, mining, forestry, and fisheries. Figure 2 exhibits the current number of active VSS in different sectors based on the ITC Standards Map.[4] According to Figure 2 the most frequently covered products are agricultural products, followed by processed food. In the agricultural sector, the number of VSS has risen markedly since the early 1990s. Over the last two decades, global markets have seen a growing supply and demand for food and other agricultural products that possess specific quality characteristics linked to composition, origin, production method, or terms of trade (Loconto & Dankers, 2014). Within the agriculture sector, certified areas continue to increase, according to ITC (2019), in 2016–2017 alone, the certified area for banana, cocoa, cotton, sugarcane, and tea grew at least 18%. Sugarcane expanded the fastest (+89%), while cotton grew by two-thirds (+67%). Up to 2017, cocoa dominated agricultural commodities in terms of certified area harvested with a share of global total equals 23.4%, followed by coffee (23.4%) tea (16.4%) and cotton (16.2%).[5]

Figure 2. Number of active VSS by sector



In recent years, there have been many discussions and debates on VSS with regard to trade. Although VSS are voluntary by law, they can become *de facto* mandatory when market penetration is so intense that companies aiming to participate in markets have no other option than using them (Smith 2009). Nevertheless, due to high compliance costs, small-scale producers might be excluded from value chains. Therefore, VSS are one of the governance systems that affect trade patterns and could potentially marginalize developing countries' exports from accessing global value chains and lucrative markets.

This paper presents a summary of the empirical literature on the trade impact of VSS. Although many studies theoretically analyze the impact of VSS on trade, empirical type of research has been underexplored. This paper presents the very few empirical studies in this topic, provides an overall analysis of the possible reasons behind lack of studies in this field, and outlnes a way forward for future research.

The rest of this paper is organized as follows. The following section first briefly reviews the theoretical debate on the trade impact of VSS, then presents a comprehensive analysis of the empirical literature in this field. Section 3 discusses and provides lessons learned from existing research and proposes ways forward. Section 4 concludes the study.

[4] Generated using International Trade Centre (ITC). 2020. *ITC Standards Map*. Retrieved from: https://www.sustainabilitymap.org/standards?q=eyJzZWxlY3RlZENsaWVudCI6Ik5PIEFGRklMSUFUSU9OIn0%3D

[5] This is based on 14 major sustainability standards that are: 4C – Common Code for the Coffee Community, Better Cotton Initiative (BCI), Bonsucro, Cotton made in Africa (CmiA), GLOBALG.A.P., Fairtrade International, Forest Stewardship Council (FSC), IFOAM – Organics International, the Programme for the Endorsement of Forest Certification (PEFC), ProTerra Foundation, Rainforest Alliance, the Roundtable on Sustainable Palm Oil (RSPO), the Round Table on Responsible Soy (RTRS) and UTZ.

2. Literature review

Discussions on VSS impacts on trade are mostly around their effect on market access and development, their legal status in the WTO, their implications for global governance, and the means of maximizing their benefits and minimizing their risks. Unfortunately, there has not been much progress achieved so far (Thorstensen et al., 2015).

This section presents in short, the theoretical debate on the trade impact of VSS, followed by a detailed review of the empirical research on this topic.

2.1 Trade impact of VSS: Theoretical perspective

There are different channels through which VSS can potentially affect trade. UNFSS (2018) shows that VSS affect trade through their effect on the structure of the market, and global value chain participation and structure. Sexsmith & Potts (2009) highlight three channels through which sustainability standards can principally change value chain structures. First, by converting products to markets that have higher demand; where consumers are willing to pay a premium price for items that have been produced under sustainable conditions. Second, by changing the relations of participants with other chain actors, policymakers, and other organizations. And third, by altering the rules of value chain involvement and participation, and the distribution of authority to make these rules.

According to the literature, international standards can be catalysts or barriers to trade. One body of the literature suggests that these standards could lead to increased exports, as VSS provide a competitive advantage to complying producers, affirm high product quality, and signal sustainable production practices that facilitate their market access to foreign markets. Masood & Brümmer (2014) define a favourable trade impact of VSS as a result of the demand enhancing effect that takes place due to the safety and quality of certified products, product differentiation, and harmonization. Besides, standards could be trade-enhancing as they reduce information asymmetries and transaction costs (Henson & Jaffee, 2008; Jaffee, 2003), and modernize the food supply chains through innovation and upgrading (Swinnen, 2007). Moreover, certification can bring benefits in the form of higher productivity and lower input costs (Graffham et al., 2007).

From the importers' point of view, according to Andersson (2019), certification to VSS could reduce variable cost by providing information and a form of quality insurance. It also reduces sunk trade costs through reduced search costs and possibly quality control costs. This is expected to positively affect both the intensive and extensive margin of trade (i.e. import volumes and the probability of trade, respectively).

However, a second body of the literature suggests that the expansion and increased influence of VSS have become an increasing concern for suppliers, in particular those in low-income countries. If VSS are de facto mandatory for specific markets, small-scale producers mainly risk being excluded from export value chains due to high compliance costs and increasing monitoring costs (Unnevehr, 2000; UNCTAD, 2008). Standards also affect the international competitiveness of domestic farmers, particularly in developing countries. If producers in developing countries are competing directly with producers in developed countries, and are in general less able to implement the requirements of VSS at a given level of cost, they could lose out (UNFSS, 2013). Masood & Brümmer (2014) state that the negative effect of VSS on trade is captured through the cost effect, either the compliance cost or delaying effect and the drive out effect. Also, other nonfinancial obstacles like financial literacy are found to constrain farmers to adopt standards (Müller & Theuvsen, 2015).

Another potential obstacle that restricts low-income countries from participating in global value chains is the regulatory environment. Marx & Wouters (2015) show that VSS are mainly adopted in developed countries with robust regulatory environments. On the other hand, low-income economy countries with fragile governance structures hardly have any entities which are certified according to VSS.

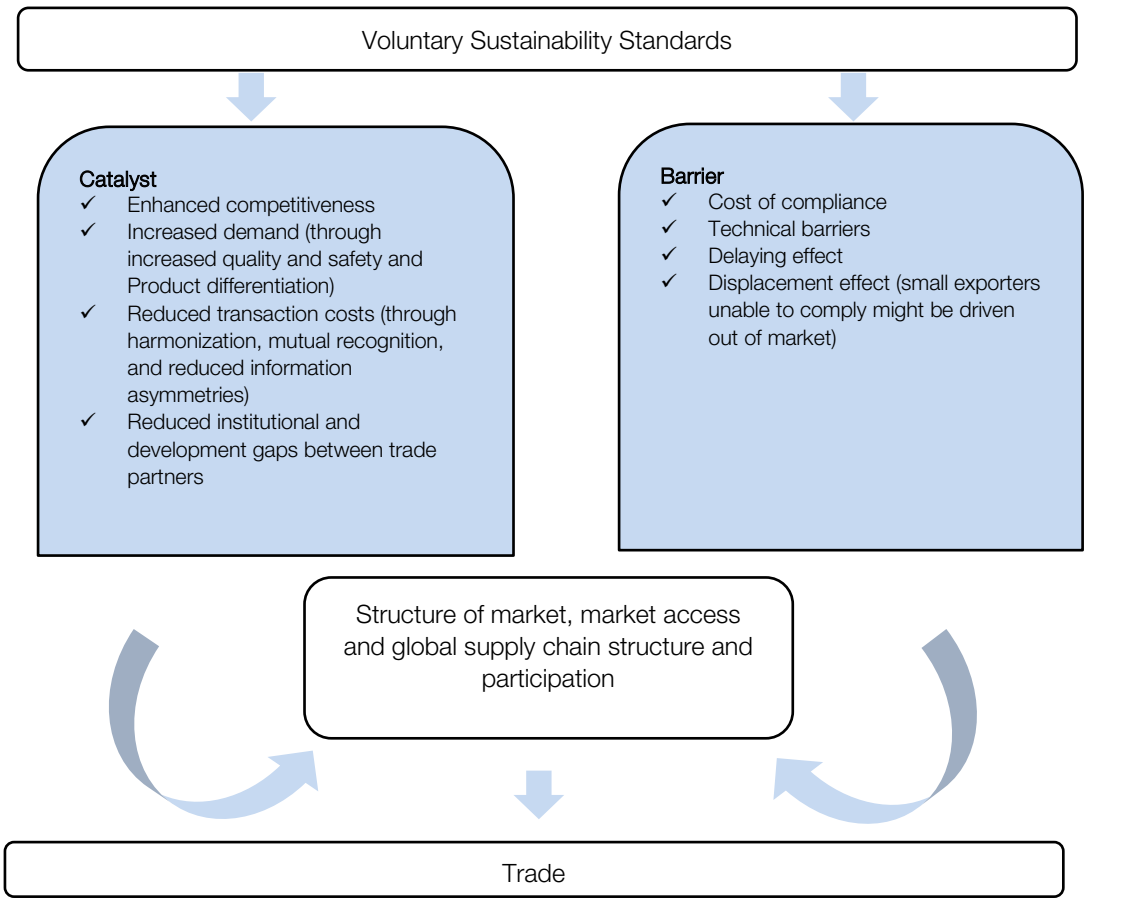
The chief argument for voluntary standards having a deleterious impact on international trade revolves around the burden of compliance costs (Hobbs, 2010). On the other hand, standards are argued to help in reducing

transaction costs between buyers and sellers and make trade more likely. Mangelsdorf (2011) and Swan (2010) conclude that standards reduce trade when the compliance costs outweigh transaction costs and foster trade vice versa. Figure 3 summarizes the theoretical literature on the trade impact of VSS and presents the channels through which VSS can potentially affect trade.

UNFSS (2018) shows that the possible impact VSS have on trade is subject to the institutional design of the standards themselves. For example, distributing the compliance costs of sustainability standards more equally among supply chain players is a way to reduce the burden on developing countries producers. Also, increasing transparency is a way to reduce transaction costs.

Most of the arguments regarding VSS impact on trade arise from theoretical perspectives and anecdotal examples. However, there is little systematic evidence for these arguments. Thus, a primary motivation for this project is simply to examine such arguments based on reviewing the empirical research on this topic. Insights from this review draw a way forward for future research by highlighting the lessons learned from previous studies and identifying the empirical gaps.

Figure 3. Channels through which VSS impact trade



2.2 Trade impact of VSS: Empirical evidence

Although international standards, regulations, and voluntary standards may, in principle, have the same trade effects (Meliado 2017), most studies deal with public standards and their impact on bilateral trade. In the food sector, Hobbs (2010) and Henson (2008) state that most of the empirical studies examining the implications of food safety and quality standards on trade between developing and developed countries are confined to public standards or codified quality meta-systems such as ISO 9000.

The focus on mandatory standards can partly be explained by data availability. Commonly, information on mandatory standards is publicly available, while data on VSS is often challenging to obtain. Beghin et al. (2015) review studies on public and private standards in the food sector and conclude that studies investigating the trade impact of private standards often suffer from a lack of representative data.

First, by reviewing keywords and abstracts, studies included in this research were selected, based on different criteria:

1. Empirical studies rather than theoretical discussions papers,
2. Focus on trade impacts rather than production level or consumer demand.
3. Focus on non-state voluntary standards rather than state voluntary standards, and
4. There is no conflict of interest; the study was not funded by a standards organization.

The search resulted in a total of nine studies (see Table 1) on a country-level and firm-level. One point to mention, these studies focused on a very narrow set of standards that are Global Good Agricultural Practice (GlobalGAP), British Retail Consortium (BRC), International Food Standard (IFS), Good Manufacturing Practices (GMP), Linking Environment and Farming (LEAF), and Good Agricultural Practice (GAP). This leaves a considerable research gap in this field.[6] The following section presents in detail the methodologies and models employed in the available empirical literature.

2.2.1 Country-level studies

Our search generated five country-level studies. All employed gravity models. The list of exporting and importing countries is presented here whenever it is available in the study.

Andersson (2019) - GloablGAP

Andersson (2019) analyzes the effect of GlobalGAP certification on 15 European Union countries' imports of 50 fresh fruit and vegetable products from 138 economies around the world.[7] The paper utilizes an annual dataset from 2009-2013 and employs a gravity model with fixed effects.

Model specifications:

- Method/Model: Ordinary least square (OLS) and Poisson pseudo-maximum likelihood (PPML) estimation of gravity Model.
- Dependent variable: The dependent variable is specified in three ways: (1) as a binary variable that takes the value 1 if an economy (i) imports product (k) from economy (j) at time (t); (2) as a logged

[6] We only include studies that focus on VSS, hence national standards and codified quality meta-systems such as ISO 9000, are beyond the scope of this review .For empirical studies on trade effect of these standards please refer to Blind et al. (2013), Masakure et al. (2009), Martincus et al. (2010), Clougherty & Grajek (2014) and for a comprehensive review refer to Swann (2010), Ronen (2017) and Timmis (2017).

[7] **Importers**: Austria, Belgium, Denmark, Finland, France, Germany, Greece, Italy, Ireland, Luxembourg, Netherlands, Portugal, Spain, Sweden and the United Kingdom of Great Britain and Northern Ireland. **Exporters:** Afghanistan, Albania, Algeria, Argentina, Australia, Austria, Bangladesh, Belarus, Belgium, Belize, Benin, Bhutan, Bolivia (Plurinational State of), Bosnia and Herzegovina, Brazil, Bulgaria, Burkina Faso, Burundi, Cambodia, Cameroon, Canada, Chad, Chile, China, Colombia, the Comoros, the Congo, Costa Rica, Côte d'Ivoire, Croatia, Cyprus, Czechia, the Democratic Republic of the Congo, Djibouti, Dominica, the Dominican Republic, Ecuador, Egypt, Estonia, Eswatini, Ethiopia, Finland, France, the Gambia, Georgia, Germany, Ghana, Greece, Grenada, Guatemala, Guinea, Guinea-Bissau, Haiti, Honduras, Hungary, Iceland, India, Indonesia, Iran (Islamic Republic of), Ireland, Israel, Italy, Jamaica, Japan, Jordan, Kazakhstan, Kenya, the Lao People's Democratic Republic, Latvia, Lebanon, Liberia, Lithuania, Luxembourg, Madagascar, Malawi, Malaysia, Mali, Malta, Mauritania, Mauritius, Mexico, Morocco, Mozambique, Namibia, Nepal, Netherlands, New Zealand, Nicaragua, the Niger, Nigeria, North Macedonia, Norway, Oman, Pakistan, Panama, Paraguay, Peru, the Philippines, Poland, Portugal, the Republic of Korea, the Republic of Moldova, Romania, the Russian Federation, Rwanda, Saint Vincent and the Grenadines, Saudi Arabia, Senegal, Sierra Leone, Singapore, Slovakia, Slovenia, South Africa, Spain, Sri Lanka, the Sudan, Suriname, Sweden, Switzerland, Thailand, Timor-Leste, Tunisia, Turkey, Uganda, Ukraine, the United Arab Emirates, United Kingdom of Great Britain and Northern Ireland; United Republic of Tanzania, United States of America, Uruguay, Uzbekistan, Viet Nam, Yemen, Zambia, Zimbabwe; State of Palestine, Hong Kong (China).

value of economy (i)'s imports of product (k) from economy (j) at time (t), and (3) as the logged weight of economy (i)'s imports of product (k) from economy (j) at time (t).

- Independent variables:
  - ✓ To measure GlobalGAP certification coverage, the logged number of certified GlobalGAP producers is employed (In an alternative specification, the logged number of certificate holders is applied).
  - ✓ Traditional gravity model variables that are: the logarithm of importers' and exporters' GDP (which are proxies for the economic mass and measure the trade potential between the trading partners); the European Union tariff at two-digit level; a dummy variable that equals 1 if the exporter has a Regional Trade Agreement (RTA) with the importing EU15 and zero otherwise; fixed effects variables that are: the importer-exporter fixed effect, a product level fixed effect, a time fixed effect.
  - ✓ Two multilateral resistance (MR) effects that are constructed by a GDP weighted average of importing and exporting economies logged distance to its trade partners.
  - ✓ The production level of the traded product in the exporting economy.

It is found that certification to GlobalGAP has a positive effect on both the probability of trading and import intensity. The result indicates that a 1 per cent increase in the number of certified producers producing a particular product is related to a 0.06 per cent increase in the probability of importing this product to EU15, and a 0.58–0.59 increase in imports. It is highlighted that the positive trade effect of GlobalGAP certification coverage holds for both high- and low-income economies. However, the positive impact is estimated to be higher for low-income than high-income economies.

Masood & Brümmer (2014) - GlobalGAP

Masood & Brümmer (2014) test the link between the intensity of GlobalGAP certification in 29 countries and banana imports to all European Union countries.[8] They employ a gravity modeling framework using disaggregated data on banana imports for the period 2010 to 2012.

Model specification:

- Method/Model: OLS estimation of gravity model
- Dependent variable: Banana imports value from country (i) to country (j) at time (t)
- Independent variables:
  - ✓ The number of banana producers under GlobalGAP is used (In an alternative specification, the number of harvested hectares under GlobalGAP certification is utilized).[9]
  - ✓ The standard gravity model variables for distance, colonial ties, regional trade agreement, common language, and time fixed effects variables.
  - ✓ Banana output of the exporting countries.
  - ✓ Banana consumption of the importing countries.
  - ✓ Lagged values of the number of banana producers or number of harvested hectares under GlobalGAP (to address the reverse causality of banana trade of GlobalGAP adoption and diffusion).

They find that the intensity of certification is associated with a higher value of banana imports to European Union. For example, it is found that a 1 per cent change in number of certified producers in exporting countries increases banana imports' value by 0.2 to 0.3 per cent. In addition, a 1 per cent change in harvested hectares

[8] **Exporters:** Brazil, Chile, Côte d'Ivoire, Cameroon, Colombia, Costa Rica, the Dominican Republic, Ecuador, Egypt, Eswatini, Ethiopia, France, Ghana, Guatemala, Honduras, India, Kenya, Lebanon, Morocco, Mexico, Mozambique, Panama, Peru, the Philippines, Saint Lucia, South Africa, Spain, Suriname and Thailand.

[9] Data availability in terms of number of hectares of certified banana is restricted, due to GlobalGAP data privacy policy, to only 14 countries.

area increases import value by around 0.2 per cent. This result is robust across all model specifications at aggregated and disaggregated levels of European Union countries.

The result also indicates that the magnitude of GlobalGAP certification measures does not differ much for top European Union importers and all European Union importers, i.e., that impact of GlobalGAP has been affecting banana imports of all European Union countries in a similar fashion.

Fiankor et al. (2020)– GlobalGAP

Fiankor et al. (2020) are the first to consider global bilateral trade flows from all producing to all importing countries. They focus on apple, banana, and grapes. Their dataset covers the period 2010–2015 and includes exports from 163 producing countries to 157 importing countries (91 apple-producing countries, 108 banana-producing countries, and 88 grape-producing countries).

Model specifications:

- Method/Model: PPML estimation of gravity model
- Dependent variable: Trade flows value of product (k) from exporting country (i) to importing country (j) in year (t).
- Independent variables:
  - ✓ One-year lag of the number of certified producers, or the share of certified land area in the total harvest area,[10] are used.
  - ✓ 'no certification' dummy that takes the value of 1 when the country has no GlobalGAP certification and 0 otherwise.
  - ✓ Instead of using GDP as a proxy for the mass of exporting countries they consider sector-specific production as a better measure of the supply-side capacity.
  - ✓ Traditional gravity model variables of distance, product-specific ad valorem tariff, regional trade agreement, dummies for sharing a common language, colonial ties, common border, the exporter-product-time, and importer-product-time fixed effects.
  - ✓ To account for possible product heterogeneity, they estimate a second model specification in which dummies for the different products (i.e. apples, bananas and grapes) interact with the lagged GlobalGAP certification variable.

The result confirms certification is a catalyst to trade, however, the trade-enhancing effect varies across products and destination markets. The estimated elasticity 1 per cent increase in (1) the number of certified producers and (2) share of GlobalGAP certified area increase bilateral trade by 0.33 and 0.45 per cent respectively. Therefore, the trade effect is driven more by growth in the area of certified farms than by new certified producers. Also, the 'no certification' variable indicates non-certified producing countries have, on average, lower exports compared to their certified counterparts. All these results remain the same regardless of the development status of the exporting country.

Fiankor et al. (2019) - GlobalGAP in the context of institutional gaps:

This paper tests the role of VSS in bridging institutional gaps between trading partners. When the quality of institutions differs widely between two countries, the authors argue that standards can act as surrogate governance institutions. They estimate a structural gravity model on a sample of European Union and European Free Trade Association (EFTA) imports of apples, bananas, and grapes from 134 economies between 2010

[10] The authors use lagged values to deal with potential reverse causality bias between certification and trade effect.

and 2015.[11] To measure the quality of institutions, they use a composite index of the six dimensions of the Worldwide Governance Indicators (WGI), which they refer to as ''governance distance''.[12]

Model specifications:

- Method/Model: OLS and PPML estimation of gravity model
- Dependent variable: Export value of product (k) from exporting economy (i) to importing economy (j) at time (t).
- Independent variables:
  - ✓ Governance distance which is a composite index of time-varying economy-pair differences in the six dimensions of the World Governance Indicators (WGI),
  - ✓ A dummy variable as a measure of the certification status of the exporting economy (in another specification, the number of certified producers in each exporting economy is used),
  - ✓ The interaction of the two variables above.
  - ✓ Traditional gravity variables that are identical to the variables used in Fiankor et al. (2020).

The result confirms a trade impeding effect of governance distance on exports, mitigated by the interaction between governance distance and GlobalGAP certification. For example, the trade reducing effect of a one standard deviation increase in the governance distance measure is 124% for non-certified economies but decreases to about 51% for certified producing economies.

The findings are robust to the product-specific analysis of apples, bananas, and grapes but also the aggregate fruits and vegetable sector, and individual six components of governance distance index when disaggregated.

The findings have important policy implications for producers and firms targeting high-value markets, but are located in economies with low quality of existing domestic public institutions. Getting certified to a standard that is accepted in the importing economy can help overcome the negative reputation effects associated with their geographical locations.

[11] **Importers:** Austria, Belgium, Bulgaria, Cyprus, Czechia, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Latvia, Liechtenstein, Lithuania, Luxembourg, Malta, Netherlands, Norway, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Switzerland, Sweden, United Kingdom of Great Britain and Northern Ireland. **Exporters:** Afghanistan, Angola, Albania, Algeria, Argentina, Armenia, Australia, Azerbaijan, Bahamas, Bahrain, Bangladesh, Barbados, Belarus, Belize, Benin, Bermuda, Bhutan, Bolivia (Plurinational State of), Bosnia and Herzegovina, Brazil, Burundi, Cambodia, Cameroon, Canada, Cabo Verde, the Central African Republic, Chile, China, Colombia, the Comoros, the Congo, Costa Rica, Cote d'Ivoire, Croatia, Cuba, the Democratic People's Republic of Korea, the Democratic Republic of the Congo, Dominica, the Dominican Republic, Ecuador, Egypt, El Salvador, Eswatini, Ethiopia, Equatorial Guinea, Fiji, Gabon, Georgia, Ghana, Grenada, Guatemala, Guinea, Guyana, Haiti, Honduras, India, Indonesia, Iran (Islamic Republic of), Iraq, Israel, Jamaica, Japan, Jordan, Kazakhstan, Kenya, Kiribati, Kuwait, Kyrgyzstan, Lao People's Democratic Republic, Lebanon, Libya, Madagascar, Malawi, Malaysia, Mali, Mauritius, Mexico, Morocco, Mozambique, Namibia, Nepal, New Zealand, Nicaragua, North Macedonia, Oman, Pakistan, Panama, Papua New Guinea, Paraguay, Peru, the Philippines, Qatar, the Republic of Korea, the Republic of Moldova, the Russian Federation, Rwanda, Saint Lucia, Saint Vincent and the Grenadines, Samoa, Saudi Arabia, Senegal, Serbia, Seychelles, Sierra Leone, Somalia, South Africa, the Sudan, Suriname, the Syrian Arab Republic, Tajikistan, Thailand, Togo, Tonga, Trinidad and Tobago, Tunisia, Turkey, Turkmenistan, Uganda, Ukraine, the United Arab Emirates, United Republic of Tanzania, United States of America, Uruguay, Venezuela (Bolivarian Republic of), Viet Nam, Yemen, Zambia, Zimbabwe; British Virgin Islands, Cook Islands, French Polynesia, Montserrat, New Caledonia.

[12] Components of Worldwide Governance Indicators: Voice and Accountability, Government Effectiveness, Control of Corruption, Regulatory Quality, Political Stability, Rule of Law.

Ehrich & Mangelsdorf (2018) – IFS

Ehrich & Mangelsdorf (2018)[13] use data of 12000 food processing firms that are certified with the IFS from 87 economies for the period 2008 to 2013.[14] The study focuses on seven sectors that are egg products, meat, fruits and vegetables, bakery products, dairy products, and beverages.

The authors estimate gravity models at three different levels: the aggregate level, by income group of the exporting economy, and at the sectoral level.

Model specification:

- Method/Model: PPML estimation of gravity model
- Dependent variable: Bilateral trade value.
- Independent variables:
  - ✓ The number of IFS certifications in the exporting economy.
  - ✓ Traditional gravity model variables (GDPs of exporter and importer economies, distance, trade agreements, tariff, common language, colonial links and contiguity).
  - ✓ Instead of estimating economy-year fixed effects, which, as stated by the authors, is computationally difficult in the case of this study due to the large data set, the Baier–Bergstrand technique is preferred and employed.

The result indicates that a 1 per cent increase in the number of certified firms increases exports on average by about 0.147 per cent to 0.27 per cent. However, this remains robust only for high- and middle-income economies and disappears for low-income economies. This result leads the authors to argue that food standards are not a suitable development tool to integrate low-income economies into high-value chains per se.

In summary, noticeably, all the country-level studies in the previous section employ gravity modeling. This is not surprising; in general, econometric approaches for modeling trade flows have focused on the gravity model specification. Moreover, gravity models are the standard approach to estimate the impact of standards on trade (Ehrich & Mangelsdorf, 2018). This is mainly because of its robust performance and limited need for parameter assumptions (Ronen, 2017).

2.2.2 Firm-level studies

Firm-level studies create insights that are complementary to evidence from a macroeconomic perspective. The search generated four studies on a firm-level.

Henson et al. (2011) – GlobalGAP

Henson et al. (2011) investigate the returns of GlobalGAP certification in terms of expansion of export sales revenue. They conduct a survey of 102 fresh produce export firms in 10 Sub-Saharan countries to estimate

[13] Although the authors use firm-level data, they still employ a gravity model and use the number of firms at country-level as an independent variable. Thus, their study fits in country-level studies.

[14] Importing and Exporting economies: **Europe:** Albania, Austria, Belarus, Belgium, Bosnia and Herzegovina, Bulgaria, Croatia, Cyprus, Czechia, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania, Luxembourg, Netherlands, North Macedonia, Norway, Poland, Romania, San Marino, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey, Ukraine, United Kingdom of Great Britain and Northern Ireland; Faroe Islands. **Africa:** Cote d'Ivoire, Egypt, Ghana, Kenya, Madagascar, Mauritius, Morocco, Namibia, Senegal, Seychelles, South Africa, Tunisia, Uganda. **America:** Antigua and Barbuda, Argentina, Brazil, Canada, Chile, Colombia, Costa Rica, Ecuador, Guatemala, Guyana, Honduras, Mexico, Nicaragua, Peru, Suriname, United States of America, Uruguay. **Asia:** Armenia, Australia, Azerbaijan, Bangladesh, China, India, Indonesia, Iran (Islamic Republic of), Israel, Kazakhstan, Malaysia, Pakistan, Papua New Guinea, the Philippines, the Republic of Korea, the Russian Federation, Sri Lanka, Thailand, the United Arab Emirates, Viet Nam.

the causal effect of certification on firm-level export performance to the European Union over the period 2000–2006.

Model specification:

- Method/Model: Propensity score matching and difference in differences between certified and non-certified firms.
- Dependent variable: Change in export value.
- Independent variables:
  - ✓ A vector of controls for each firm.
  - ✓ A variable capturing the mean impact of certification.

The authors started their model with a vector of firm-specific and observable variables that might influence export sales performance, in addition to a dummy variable depicting whether a firm is certified. To correct for selection bias on the observable variables, they use propensity score matching methods; that each exporting firm certified to GlobalGAP, has a very similar control firm that is non-certified. [15] This procedure matches exporting firms according to a large vector of relevant characteristics. They used the nearest neighbor matching (NNM) and Kernelbased matching (KBM) estimators (Becker & Ichino, 2002). Then Difference in Differences (DID) matching estimator, which allows for time-invariant differences in the outcomes between certified and non-certified firm pairs, is used. Finally, the difference in the differences between certified and non-certified firms is applied.

It is found that firms that have achieved certification, on average, have export sales revenue that is around €2.6 million higher than it would otherwise have been. Thus, the authors conclude that the high cost of certification is a profitable investment.

In addition, the authors highlight that the high compliance costs can lead to exclusion from exported markets. Moreover, the technical and financial assistance for standards implementation are essential tools to support exporters from developing countries.[16]

Schuster & Maertens (2015) – GlobalGAP, LEAF, GAP, BRC, GMP and IFS

Schuster & Maertens (2015) analyze firm-level trade effects of certification to private standards in the fresh asparagus export sector in Peru. They use 18-years panel data from 87 firms. The standards included in their study are GlobalGAP, LEAF, GAP, BRC, GMP, and IFS.[17]

By controlling for export persistence, time-constant unobserved heterogeneity, and reverse causality, the authors made methodological improvements that resulted in more correct estimates of the impact of private standards on firms' export performance.

Model specifications:

- Method/Model: Generalized method of moments (GMM)
- Dependent variable: The dependent variable is specified in five different ways: (1) as a dummy variable equalling one if a firm is exporting; (2) as a logarithm of the export volume or (3) export value

[15] While certification may enhance export performance, it is also possible that export performance leads to certification. The assumption is that the researchers had access to some variables correlated with both export performance and certification to GlobalGAP such that, conditional on these variables, assignment to certification is random.

[16] They include a dummy variable (Assistance) in the model. It takes the value of one if the firm had received any technical and/or financial assistance specifically for implementation and/or certification to GlobalGAP. This variable coefficient turns out to be positive and significant.

[17] The paper also included other standards that aren't recognized as VSS that are: Safe Quality Food Institute (SQF2000), Tesco Nurture (TESCO: supermarket standard), Hazard Analysis and Critical Control Point (HACCP), Safe Quality Food Institute (SQF1000).

being positive when the firm exports or zero when the firm does not export; and (4) as a logarithm of the export volume or (5) the export value, conditional on exporting in that year

- Independent variables:
  - ✓ A dummy variable equalling one if the firm is certified to any type of standards
  - ✓ A vector of a large set of observable firm characteristics (the type and size of the firm, experience, access to foreign capital, tax regime, management changes, and location)
  - ✓ A dummy to control for common macroeconomic effects.
  - ✓ A one-year lag of the respective dependent variable to account for export persistence over time.
  - ✓ They also decompose the error term in a time-constant and a time-varying component to consider the role of unobserved firm heterogeneity.

No evidence is found that certification to private standards in general and to specific individual private standards affects firms export performance to different markets, neither at the extensive margin nor at the intensive margin, and neither on export volumes nor on export values.

The results indicate that exports are sticky and that unobserved firm characteristics play an essential role in determining both export performance and certification decisions. One point to stress is although the results imply that private standards do not enhance trade, it does not necessarily mean that they impede trade either.

Latouche & Chevassus-Lozza, (2015) – IFS and BRC

Latouche & Chevassus-Lozza (2015) study whether certified firms are more export-oriented than non-certified firms by investigating the impact of BRC and IFS certification on French agri-food firms exports to all European Union markets. Using a data set of 2,942 agri-food firms, exporting or not, including 573 firms that are IFS and/or BRC certified in 2007.

They generate the productivity threshold to export and test whether this threshold is significantly lower for certified firms than for non-certified ones.[18]

Model specifications:

- Method/Method: Maximum likelihood
- Dependent variable: Productivity threshold to export[19]
- Independent variables:
  - ✓ Importer's market size (derived as the share of country (j) in total EU25 imports of subsectors).
  - ✓ Distance from the head office of exporting firms to the capital of the importing countries.
  - ✓ An indicator of the potential supply in importing countries.[20]
  - ✓ Subsector fixed effects.
  - ✓ Importing country fixed effects.

The empirical strategy consists of estimating the threshold for different samples of certified and non-certified firms. To control for the role of productivity and size in the estimation, they use the propensity score matching method to ensure that the effect of certification on export behaviour is disentangled from the effect of productivity. They assume that the firm's productivity follows a Pareto distribution, then estimate the probability that the productivity of a non-certified firm is above the threshold to enter a specific market using the maximum

[18] A lower value of the threshold indicates an increase in the probability for French firms to export.

[19] Chaney (2008) shows that there is a productivity threshold, specific to each foreign market, above which firms with a higher productivity may enter the market. In Chaney (2008), this threshold is unique to each market. In the present paper, they assume that at entry to a given market, two thresholds can coexist: one for certified firms, another for non-certified firms.

[20] This is derived from the multilateral resistance index from Anderson and van Wincoop (2004). This index accounts for the potential supply available in country j and is corrected for proximity (in distance, language and common borders) between j and its potential partner.

likelihood method. As country fixed effects cannot be identified directly in the specified model, and to solve this identification problem, they account for the nine agri-food subsectors to which the firms belong. Consequently, the demand addressed by importing countries to these subsectors, as well as the potential supply of these subsectors on importing markets are addressed.

BRC-certified firms are found to be more export-oriented than non-certified firms. However, no export effect is found for IFS certified firms. Also, when developing and estimating Chaney's model, they find that the firms certified with BRC face lower trade costs to access certain European countries.

Colen et al. (2012)– GlobalGAP

Colen et al. (2012) compare the export volumes, market shares, and length of the export season of mango (n=72) and green beans (n=196) certified and non-certified Senegalese companies between 2005 and 2010.

It is found that certified companies have higher market shares and larger export volumes to the European Union both for beans and mangoes. However, the difference in growth rates of exports between certified and non-certified mango companies is not statistically significant. On the other hand, although beans exports experienced a decline over time, the exports of certified companies decreased less in percentage terms than those of non-certified companies, with the difference marginally statistically significant. As for the length of the export season, a significant difference between certified and non-certified companies is found for green beans, but not for mangoes.[21]

The previous section summarizes empirical studies on the trade impact of VSS. It is also important to highlight that other studies that focus on assessing the effect of certification on domestic local sales do exist. As our primary concern is the international trade impact of VSS, these studies are beyond the scope of the current review.[22]

[21] No model specification is presented in the original paper. The authors however compare between the export volumes of certified and non-certified firms by running a one side t-test to check the significance of the difference in exports.

[22] For interested readers, refer to Subervie and Vagneron (2013) who investigate the impact of GlobalGAP certification on the local marketing performances of lychee farmers in Madagascar. As farmers don't engage in exporting activities, they sell their products to local exporters. Their results suggest a positive average impact of certification on both the quantities sold and the prices received by certified producers.

Table 1. Summary of the reviewed papers

Level	Authors	Number of exporting economies/firms	Modeled variable	Standard	Product and Sector	Standards Data Source (as mentioned in the study)	Technique/Model	Result: Trade impact of VSS
Country-level	Andersson (2019)	138 economies	Import value to EU15	GlobalGAP	fresh fruits and vegetables	GlobalGAP's data mining and statistics	OLS and PPML estimation of gravity model	Favourable
	Masood & Brümmer (2014)	74 countries	Import value to the European Union	GlobalGAP	Banana	GlobalGAP Head office Cologne	OLS estimation of gravity model	Favourable
	Fiankor et al. (2019)	134 economies	Export value to the European Union/EFTA	GlobalGAP	Apple, banana and grapes	Not specified	OLS and PPML estimation of gravity model	Favourable,
	Fiankor et al. (2020)	Global bilateral trade	Bilateral trade value	GlobalGAP	Apple, banana and grapes	GlobalGAP Head office Cologne	PPML estimation of gravity model	Favourable
	Ehrich & Mangelsdorf (2018)	Firms from 87 economies	Bilateral trade value	IFS	Seven manufactured food products: egg products, meat, fruits and vegetables, bakery products, dairy products, and beverages	IFS-auditing database	PPML estimation of gravity model	Favourable (however, remains robust only for high- and middle-income countries)
Firm -level	Henson et al. (2011)	102 firms in 10 sub-Saharan countries	Exports value to the European Union	GlobalGAP	Fresh produce	Surveys	Difference in Difference (OLS) + propensity score matching	Favourable
	Schuster & Maertens (2015)	87 Peruvian firms	Exports volume and value to all trading partners	GlobalGAP, LEAF, GAP, BRC, GMP, and IFS	Asparagus	Primary data from a survey among a representative sample of export companies.	GMM	No effect
	Latouche & Chevassus-Lozza, (2015)	2942 French firms	Productivity threshold to export to European markets	IFS BRC	Agri-food	The official websites of the two standards IFS and BRC	Maximum Likelihood	Mixed effect
	Colen et al. (2012)	72 mango exporting firms and 196 green beanexporting firms in Senegal	Export volume to European Union	GlobalGAP	Mango and bean	Survey and interviews	Compared export volumes, market shares and length of the export season between certified and non-certified firms.	Mixed effect

3. Lessons and way forward

3.1 Discussion and Results

The summary presented above does not provide a convincing answer to whether VSS have an impact on trade nor to the type (direction) of impact. While country-level studies (Andersson, 2019; Masood & Brümmer, 2014; Ehrich & Mangelsdorf; and Fiankor et al., 2019) conclude that certifications have a favorable impact on trade, firm-level studies find no significant trade enhancing effect (Schuster & Maertens, 2015), or a mixed effect. For example, Latouche & Chevassus-Lozza (2015) find at most, one of the analyzed certification schemes in their paper to have a favorable impact on export. Also, Colen et al. (2012) conclude GlobalGAP certificate has a significant beneficial impact on the exports of one product out of the two they investigated. These mixed results represent a huge challenge when trying to derive policy recommendations for developing countries and VSS. Furthermore, the studies do not throw additional light on other development parameters, such as income distribution, wage increase or poverty reduction.

One point to stress it that with one exception, the country-level studies test the impact of the same standard (that is, GlobalGAP). Moreover, although the studies test different products, all employed the number of certified producers as a measure of the certification intensity. Also, three of the country-level studies focus on imports to European Union countries (i.e. a similar destination). Another point to mention, all the five studies employed gravity modelling framework.

Although some of the studies generate similar conclusions on the aggregated level, conflicted results are obtained on a disaggregated level. For example, when countries were disaggregated based on their income level, Andersson (2019) finds that the positive trade effect of increasing certification coverage is higher for low-income than high-income economies. While Ehrich & Mangelsdorf (2018) find that the trade enhancing effect of certification remains robust only for high and middle-income economies and disappears for low-income economies. On the other hand, Fiankor et al. (2020) find the positive VSS impact on trade to remain the same regardless of the development status of the exporting economy.

While most of the studies use a very similar methodology by employing the inventory approach, i.e. using counts of certifications/certified producer..etc. as measures of VSS coverage, and examine their link to trade, Fiankor et al. (2019) employ a different approach by testing the role of VSS in bridging the institutional gaps between trading partners. This helps in understanding one potential channel through which VSS impact trade. The research findings have significant policy implications in particular for developing countries. We encourage researchers to pay more considerable attention to assessing the possible channels and means through which VSS could potentially impact trade.

Although firm-level analysis is essential in testing firms' specific characteristics impact on VSS adoption and its implications for trade, firm-level studies, especially the ones that focus on a single country, generate very case-specific conclusions.

To summarize:

1. The small number and recency of studies in this field highlights that this area is indeed underexplored by researchers and signifies the need to dig more into it.
2. The available evidence of the trade impact of VSS is relatively weak and case-specific for many reasons:
    - Studies are extremely few in number.
    - The studies focus on very few standards (GlobalGAP, IFS, BRC, LEAF, GMP). According to ITC Standards Map, there are 264 active VSS (up to May 2020). Thus, it is not feasible to draw a conclusion based on a small sub-sample of standards. This narrow focus leads to gaps in the literature.

- The studies focus only on the agricultural sector. Although two studies included a wide range of agricultural products (Andersson, 2019; Ehrich & Mangelsdorf, 2018), the evidence remains limited to one sector. Also, most of the studies focus on a single product rather than a range of products' HS codes. The focus on agricultural products is justifiable since the most frequent VSS covered products are agricultural. However, a variety of other sectors where VSS also play a substantial role is overlooked.
- Most of the available studies focus on the same trading partner that is the European Union, either all European Union countries or a sub-sample.

3. Many questions have not been adequately addressed yet. Among these questions: (1) Are developing countries more affected by VSS? (2) Are small-medium sized firms more adversely affected by VSS? (3) How are particular sectors/products affected by VSS? And (4) Do Harmonization and Mutual Recognition (for example, in the case of organic certification) necessarily impact trade positively?
4. From a policy point of view, additional questions need to be addressed, for example, whether policies that support VSS adoption (directly or indirectly) ultimately have an effect on VSS adoption and therefore trade? What policies generated an observed impact on trade? What VSS related policies work the best for developing countries in their attempt to compete in the global market?
5. Possible reasons behind the lack of empirical evidence are:
    - VSS are voluntary, which makes it less tractable by empirical trade researchers. However, the demand pattern and current trend suggest that VSS have become a de facto mandatory. The increasing policy interest and the debate on VSS, along with the lack of empirical evidence, make this a very promising research area that would contribute to both researchers and policymakers.
    - Based on our research on the publicly available statistical data on VSS, the main challenge in the quantification of VSS is data availability. In the following sections, the data challenge is highlighted.

3.2 Issues and Challenges

3.2.1 Measures of VSS coverage

The first challenge is the measures of VSS coverage. The literature employs mainly two variables as a measure of VSS coverage in the exporting countries that are:

1. The number of certified firms or producers and certificate holders (i.e. number of certifications). The difference between certified producers and certificate holders allows to highlight group certified producers. In the case of group certification, it is the group that becomes the certificate holder, not the individual producers.
2. Number of hectares harvested under the specific standard/ share of certified land area in total harvest area.

The literature uses the inventory approach i.e. a simple count of the number of certifications as a measure of VSS coverage and intensity. This approach contains several drawbacks, including that it doesn't differentiate between all existing standards as it gives equal weight to all of them. Also, it doesn't distinguish between small and large firms (when using firm-level data). Other factors that consider the size of the firm and can also be utilized if available include certified area and certified production volume.

An additional challenge is that it is hard to identify the specific channels through which VSS can potentially impact trade. As the link between trade and VSS can be looked at through different channels and factors, to name a few:

1. Product-related factors: for example, price premium, enhanced quality, and product differentiation.

2. Certification-related factors: for example, the cost of compliance, and technical barriers.
3. Trust-related factors: for example, reduced institutional gaps and narrowed development gaps between trade partners.

Measures/proxies of these factors can be considered and modeled. In addition, the interactions between these variables and VSS coverage/intensity variables (inventory approach variables) can also be considered.

3.2.2 The data challenge

Clougherty & Grajek (2014) mention three challenges that have held back the empirical scholarship on the standards-trade relationship, which are measurement, varied effects, and endogeneity. We add to these challenges the absence of publicly available data on standards. The lack of data means that the existing literature mainly consists of case studies focusing on specific products, countries, or producer groups. To our knowledge, the sources that provide publicly available data include:

1. FiBL Statistics

FiBL in a cooperation with the International Federation of Organic Agriculture Movements (IFOAM- Organics International) provides data on organic certification on an aggregated level (all organic crops) and per selected 126 crops.[23] The data is available per country, regions (Europe, Africa, Asia, Latin America, Northern America Oceania, and World), and for special groups (EFTA and DAC- countries on the list of the OECD's Development Assistance Committee), for variables listed in Table 2.[24]

Table 2. Variables available in FiBL Statistics

Aggregated Level (all crops)	Per selected crops level
• Organic area [ha] • Organic area share of total farmland [%] • Number of organic exporters • Number of organic importers • Number of organic processors • Number of organic producers • Organic exports [Million €] • Organic imports [Million €] • Organic retail sales [Million €] • Organic per capita consumption [€/person] • Organic retail sales [Million €] • Organic retail sales growth (1 year) [%] • Organic retail sales share [%]	• Organic area [ha] • Organic area fully converted [ha] • Organic area share [%] • Organic area under conversion [ha]

[23] **The selected crops are:** Almonds, Andean grains, Apples, Apricots, Avocados, Bananas, Barley, Beans, Berries, Berries, (other), Black chokeberries, Blackberries, Blueberries, Brassicas, Buckthorn, Buckwheat, Cherries, Chestnuts, Cocoa, Coconuts for oil, Coconuts eating, Coconuts (no details), Coconuts (other), Coffee arabica, Coffee associated with (other) crops, Coffee robusta, Coffee (no details), Coffee (other), Cotton, Cranberries, Currants, Dates, Dry pulses mixed, Easy peelers, Egg fruit, Elder, Figs, Flax, Fodder beet, Fresh herbs, Fresh vegetables and melons, Fruit, temperate (no details), Fruit temperate, (other), Fruit, tropical and subtropical (no details), Fruit, tropical and subtropical (other), Gooseberries, Grain maize and corn cob mix, Grapefruit/Pomelos, Grapes (no details), Grapes (other), Grapes raisins, Grapes table, Grapes wine, Hazelnuts, Hemp, Hops, Industrial crops (no details), Industrial crops (other), Kiwis, Lemons and limes, Lentils, Linseed (oil flax), Litchi, Lupine, Mustard, Nectarines, Nuts (no details), Nuts (other), Oats, Oil palm, Oilseeds (no details), Oilseeds (other), Olives (no details), Olives oil, Olives (other), Olives table, Opuntia, Oranges, (other) cereals, (other) fodder roots, Peaches, Peaches and nectarines (no details), Pears, Peas, Persimmons, Pineapples, Pistachios, Plums, Pome fruit (no details), Pome fruit (other), Pomegranate, Potatoes, Protein crops (no details), Protein crops (other), Pulses, Pumpkin seeds, Quinces, Rape and turnip rape, Raspberries, Rice, Root crops (no details), Root crops (other) , Rosehips, Rye, Soybeans, Spelt, Stone fruit (no details), Stone fruit (other), Strawberries, Sugar beet, Sugarcane, Sunflower seed, Tangerine, Tea, Textile crops, (no details) Textile crops (other),Tobacco, Triticale, Vegetables and fruit, Vegetables, fruit, Vegetables, leafy or stalked, Vegetables (no details), Vegetables (other), Vegetables, root tuber and bulb, Walnuts, with shell,Wheat.

[24] Data available at https://statistics.fibl.org/

Although the most widely adopted certification in the agricultural sector is organic certification, we didn't find any study on the trade impact of organic certification. One challenge when modelling organic data remains the comparability of different countries' data. However, given that demand for organic remains concentrated in North America and Europe (FiBL, 2020), and that the United States, European Union, Switzerland, Canada have entered a mutual recognition of organic (FiBL, 2020), it seems reasonable to assume no differences exist in organic recognition between countries. Also, organic standards recognized by IFOAM are, up to a point, harmonized.

2. ITC Standards Map

ITC Standards Map provides data on the number of standards available in different sectors, products, regions, and countries (both producing and destination region/country).[25] One problem is that one can obtain the number of standards at the time of search only, i.e., it doesn't provide the change/trend over time.

3. The State of Sustainable Markets: Statistics and emerging trends

ITC's report "State of Sustainable Markets" provides insight into the evolution of certified agriculture and forestry. It includes data for about 14 major sustainability standards that are: 4C – Common Code for the Coffee Community, Better Cotton Initiative (BCI), Bonsucro, Cotton made in Africa (CmiA), GLOBALG.A.P., Fairtrade International, Forest Stewardship Council (FSC), IFOAM – Organics International, the Programme for the Endorsement of Forest Certification (PEFC), ProTerra Foundation, Rainforest Alliance, the Roundtable on Sustainable Palm Oil (RSPO), the Round Table on Responsible Soy (RTRS) and UTZ. For selected products that are bananas, cocoa, coffee, cotton, oil palm, soybeans, sugarcane, tea, and forestry products. Up to now, there are four reports available (2015, 2017, 2018, and 2019).[26] The 2015 report gives disaggregated data of the years 2013 and/or 2014; the 2017 report gives data of the year 2015 and the 2018 report presents data of the year 2016; while 2019 provides data of the year 2017.

The report presents data on four variables: area harvested (or certified) (ha), the share of total area, production volume (MT), and the number of producers, for selected top countries. One challenge is that the list of variables and countries changes within the same report from a standard to another and a product to another, and across different reports (i.e. across time). This adds more challenges to the possibility of building and utilizing a large panel dataset. Another problem includes the possibility of double certification that leads to double-counting the variables.[27] Overall, the amount of missing data makes it hard to build a reliable model.

4. Standards organization

Most of the reviewed papers had access to the data through the standard organization (see Table 1).

[25] Data available at https://www.sustainabilitymap.org/standards?q=eyJzZWxlY3RlZENsaWVudCI6Ik5PIEFGRkIMSUFUSU9OIn0%3D

[26] Available at http://www.intracen.org/uploadedFiles/intracenorg/Content/Publications/collections/Standards.pdf

[27] **Multiple certification problem**: Additional challenges include the possibility of double certification that – when employing the inventory approach- leads to double counting variables that capture the VSS coverage. According to ITC (2019), in 2017, Forest Stewardship Council (FSC), and the Programme for the Endorsement of Forest Certification (PEFC), agreed to jointly collect and publish data about double certification on a country-level in response to requests from stakeholders. According to their findings, in 2018, a total of 431.4 million hectares were certified worldwide. Of this area, 16.5% was FSC- and PEFC-certified. Based on this, in their calculation of area certified, ITC (2019) adjusted their totals for multiple certifications, assuming that 15% is double-certified, based on FSC and PEFC assumptions. A similar approach can be conducted by other standards organizations in a collaboration with researchers to adjust the available VSS coverage data to multiple certifications in order to overcome the double counting challenge. When using surveys, direct questions that investigate the existence of multiple certifications within the farm/firm, the types of certification, and the timing each certification was adopted, need to be included in the survey/interview questions. Double certification also poses other potential research questions; for example, from a cost-benefit angle, what are the impacts of double certification on market access and trade in comparison to single certification.

5. Surveys

Some of the reviewed studies collected data using surveys and interviews with individual firms, then based their analysis on a comparison of the export performance and revenues between certified and non-certified samples. Survey-based analysis tends to be a firm-level analysis, and – unless performed over a considerable landscape across countries – generates cases specific conclusions.

One way to overcome these drawbacks is to establish a research collaboration/competition, where interested researchers from all over the world follow a similar methodology in data gathering and analysis. Hence, utilizing the same questionnaire and interview questions in a way that enables merging these data into one model/analysis technique to arrive at conclusive derivations.

In summary, there is a need for higher transparency in terms of data availability from all stakeholders.

4. Conclusion

The current paper surveys the relevant empirical literature on the trade effects of VSS. Most of the researches that quantify the impacts of standards on trade focus on mandatory public standards. One of the main findings is a worrying lack of empirical literature on the trade impact of voluntary standards. We found only nine empirical studies that quantitatively model and assess VSS impact on trade. Five of these studies are country-level studies, while the rest are firm-level ones. Moreover, all of them are conducted within the last decade, hence are recent studies.

The focus on mandatory standards in comparison to VSS can partly be explained by data availability. Through this research, we attempt to present the publicly available sources of data. However, the number of resources and amount of data openly available might still remain insufficient in performing a comprehensive empirical analysis. Thus, we stress the significance of increasing transparency in terms of data availability.

Given the small number of available empirical studies, and since the currently available evidence is, to a great extent, case-specific, it is not possible to arrive at a conclusive result. Although, all reviewed country-level studies find the enhancing trade effect of VSS more dominant on the aggregated level, in terms of increase of exports, imports, bilateral trade level, and the probability of trading, conflicted results are obtained on the disaggregated level.

We call for more research on how VSS can potentially affect the trade opportunities between countries, with a particular focus on their impact on low and middle-income countries. As these countries are expected to be the most harmed by the discriminatory effect that VSS might impose. Also, more considerable recognition needs to be devoted to the channels through which VSS potentially impact trade.

Empirical research across countries, sectors, and products is essential to identify the policy implications related to the trade effects of VSS and highlight the key factors that determine whether VSS increase trade restrictiveness or possibly impact demand favorably.

Bibliography

Anderson J E and Wincoop E van (2004). Trade Costs. Journal of Economic Literature, 42, 3,691–751.

Andersson A (2019). The trade effect of private standards. In *European Review of Agricultural Economics*. https://doi.org/10.1093/erae/jby027.

Becker S O and Ichino A (2002). Estimation of Average Treatment Effects Based on Propensity Scores. *The Stata Journal: Promoting Communications on Statistics and Stata*. https://doi.org/10.1177/1536867x0200200403.

Beghin J C, Maertens M and Swinnen J (2015). Nontariff Measures and Standards in Trade and Global Value Chains. *Annual Review of Resource Economics*. https://doi.org/10.1146/annurev-resource-100814-124917.

Blind K, Mangelsdorf A and Wilson J S (2013). Mutual recognition of accreditation: Does it matter to trade? Evidence from the food, beverage, and tobacco industry. *Frontiers of Economics and Globalization*. https://doi.org/10.1108/S1574-8715(2013)0000012017.

Cashore B, Auld G and Newsom D (2004). Governing through markets: Forest certification and the emergence of non-state authority. In *Governing Through Markets: Forest Certification and the Emergence of Non-State Authority*.

Chaney T (2008). 'Distorted Gravity: The Intensive and Extensive Margin of International Trade',American Economic Review, 98, 4, 1707–21.

Clougherty J A and Grajek M (2014). International standards and international trade: Empirical evidence from ISO 9000 diffusion. *International Journal of Industrial Organization*. https://doi.org/10.1016/j.ijindorg.2013.07.005.

Colen L, Maertens M and Swinnen J (2012). Private Standards, Trade and Poverty: GlobalGAP and Horticultural Employment in Senegal. *World Economy*. https://doi.org/10.1111/j.1467-9701.2012.01463.x

Ehrich M and Mangelsdorf A (2018). The Role of Private Standards for Manufactured Food Exports from Developing Countries. *World Development*. https://doi.org/10.1016/j.worlddev.2017.08.004.

Fiankor D D D, Flachsbarth I, Masood A and Brümmer B (2020). Does GlobalGAP certification promote agri-food exports? *European Review of Agricultural Economics*. https://doi.org/10.1093/erae/jbz023.

Fiankor D D D, Martínez-Zarzoso I and Brümmer B (2019). Exports and governance: the role of private voluntary agrifood standards. *Agricultural Economics (United Kingdom)*. https://doi.org/10.1111/agec.12488.

FiBl (2018). The world of organic agriculture: Statistics and emerging trends. Retrieved from https://shop.fibl.org/chen/mwdownloads/download/link/id/1093/

FiBl (2020). The world of organic agriculture: Statistics and emerging trends. Retrieved from https//shop.fibl.org/CHen/mwdownloads/download/link/id/1294/?ref=1.

Graffham Ga, Karehu E and MacGregor J (2007). Impact of EurepGAP on small-scale vegetable growers in Kenya. *Fresh Insights*.

Henson S (2008). The role of public and private standards in regulating international food markets. *Journal of International Agricultural Trade and Development*.

Henson S and Jaffee S (2008). Understanding developing country strategic responses to the enhancement of food safety standards. *World Economy*. https://doi.org/10.1111/j.1467-9701.2007.01034.x

Henson S, Masakure O and Cranfield J (2011). Do Fresh Produce Exporters in Sub-Saharan Africa Benefit from GlobalGAP Certification? *World Development*. https://doi.org/10.1016/j.worlddev.2010.06.012.

ITC (International Trade Centre) (2017). Social and environmental standards: Contribution to more sustainable value chains. Geneva: ITC (http://www.intracen.org/uploadedFiles/intracenorg/Content/Publications/1ITC-EUI_Social_environmental_standards_Low-res.pdf).

ITC (International Trade center) (2019). The state of sustainable markets (2019): Statistics and emerging trends. Geneva: ITC. http://www.intracen.org/uploadedFiles/intracenorg/Content/Publications/Sustainabile%20markets%202019%20web.pdf.

Jaffee S (2003). From Challenge to Opportunity: Transforming Kenya's Fresh Vegetable Trade in the Context of Emerging Food Safety and Other Standards in Europe. *World Bank Working Paper*.

Latouche K and Chevassus-Lozza E (2015). Retailer Supply Chain and Market Access: Evidence From French Agri-food Firms Certified with Private Standards. *World Economy*. https://doi.org/10.1111/twec.12191.

Liu P (2009). Private Standards in International Trade: Issues and Opportunities. *Paper Presented at the WTO's Workshop on Environment-Related Private Standards.*

Loconto A and Dankers C (2014). Impact of Voluntary Standards on Smallholder Market Participation in Developing Countries. In *Food And Agriculture Organization Of The United Nations*.

Mangelsdorf A (2011). The role of technical standards for trade between China and the European Union. *Technology Analysis and Strategic Management*. https://doi.org/10.1080/09537325.2011.592267.

Martincus C V, Castresana S and Castagnino T (2010). ISO Standards: A Certificate to ExpExports? Firm-Level Evidence from Argentina. *Review of International Economics*. https://doi.org/10.1111/j.1467-9396.2010.00915.x

Marx A and Wouters J (2014). Competition and Cooperation in the Market of Voluntary Sustainability Standards. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2431191.

Marx A and Wouters J (2015). Is everybody on board? Voluntary sustainability standards and green restructuring. *Development (Basingstoke)*. https://doi.org/10.1057/s41301-016-0051-z.

Masakure O, Henson S and Cranfield J (2009). Standards and export performance in developing countries: Evidence from Pakistan. *Journal of International Trade and Economic Development*. https://doi.org/10.1080/09638190902986538.

Masood A and Brümmer B (2014). Impact of GlobalGAP Certification on EU Banana Imports: *GlobalFood Discussion Papers*.

Meliado F (2017). Private Standards, Trade, and Sustainable Development: Policy Options for Collective Action. ICTSD Issue Paper.

Müller A and Theuvsen L (2015). Financial literacy and food safety standards in Guatemala: The heterogeneous impact of GlobalGAP on farm income. *GlobalFood Discussion Papers*.

Public and Private Standards for Food Safety and Quality: International Trade Implications. (2010). *Estey Centre Journal of International Law and Trade Policy*.

Ronen E (2017). Quantifying the Trade Effects of NTMs: A Review of the Empirical Literature. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3010217.

Schuster M and Maertens M (2015). The impact of private food standards on developing countries' export performance: An analysis of asparagus firms in Peru. *World Development*. https://doi.org/10.1016/j.worlddev.2014.08.019.

Sexsmith K and Potts J (2009). Voluntary Sustainability Standards and Value Chain Governance: How sustainability standards affect the distribution of decision-making power in global value chains. *International Institute for Sustainable Development (IISD)*. https://doi.org/10.13140/RG.2.2.33647.48803.

Smith G (2009). Interaction of Public and Private Standards in the Food Chain. In OECD Food, Agriculture and Fisheries Working Papers, No. 15, OECD Publishing.

Subervie J and Vagneron I (2013). A Drop of Water in the Indian Ocean? The Impact of GlobalGap Certification on Lychee Farmers in Madagascar. *World Development*. https://doi.org/10.1016/j.worlddev.2013.05.002.

Swan G P (2010). International Standards and Trade: A Review of the Empirical Literature. *OECD Trade Policy Working Papers*. https://doi.org/10.1787/5kmdbg9xktwg-en.

Swinnen J F M (2007). Global supply chains, standards and the poor: How the globalization of food systems and standards affects rural development and poverty. In *Global Supply Chains, Standards and the Poor: How the Globalization of Food Systems and Standards Affects Rural Development and Poverty*. https://doi.org/10.1079/9781845931858.0000.

Thorstensen V,Weissinger R and Sun X (2015). Private Standards- Implications for Trade, Development, and Governance, E15. Geneva, International Centre for Trade and Sustainable Development (ICTSD) and World Economic Forum

Timmis H (2017). The Impact of Standards on Developing Country Exports. K4D Helpdesk Report. Brighton, UK: Institute of Development.

UNCTAD (2008). Private-sector standards and national schemes for good agriculturalpractices: implications for exports of fresh fruit and vegetables from Sub-Saharan Africa. UNCTAD/DITC/TED/2007/13. Geneva: United Nations Conference on Tradeand Development.

UNFSS (2013). 1st Flagship Report of the United Nations Forum on Sustainability Standards (UNFSS). Retrieved from https://unfss. org/home/flagship-publication/

UNFSS (2018). 3rd Flagship Report of the United Nations Forum on Sustainability Standards (UNFSS). Retrieved from https://unfss.org/home/flagship-publication/

Unnevehr L J (2000). Food safety issues and fresh food product exports from LDCs. *Agricultural Economics*. https://doi.org/10.1016/S0169-5150(00)00095-5.