Qualities of Quality Standards?
The likelihood of compliance with sustainability standards in retail

Paper to be presented during the
International Conference on Management in AgriFood Chains and Networks
Wageningen University, 28 May - 30 May 2008

Jeroen van Wijk[1]*, Myrtille Danse**, Rob van Tulder*

* RSM Erasmus University Rotterdam, The Netherlands
** Agricultural Economics Research Institute/ WUR, The Netherlands

Correspondence:
Jeroen van Wijk, Department of Business-Society Management, RSM- Erasmus University, Burgemeester Oudlaan 50, T7-16, 3062 PA Rotterdam, email: jwijk@rsm.nl

Abstract

Globalization of food sourcing and increased sustainability awareness among consumer has triggered the emergence of a plethora of voluntarily adopted quality standards addressing issues related to food safety, environmental protection, and socio-economic conditions. This paper assesses the relative qualities of these standards. Are some standards better than others? The study examines 30 standards that are relevant to global agri-food retail chains, in terms of their sustainability scope and compliance likelihood. It is found that, on average, the likelihood of compliance remains rather low. Standards initiated by NGOs or non-profit organisations have the highest compliance likelihood, followed by standards initiated by business associations and individual firms. The position in the global value chain of standard initiators proofs an important predictor of the standard's qualities. Other than expected, lower scores are found with standards initiated by partnerships. This might an indication of their relatively recent inauguration. Product-specific standards have higher compliance likelihood than generic standards.

Keywords: global value chain governance, sustainability, benchmarking, compliance likelihood, quality standards

[1] **Acknowledgement:** The authors want to thank Wouter de Bruijne, Jorrit Reintjes, and Vincent van Bruchem for their valuable support during the research process.

1. Introduction

There is a growing interest in the development of common strategies to improve the sustainability of production and trade of (tropical) food sourced in developing countries. Requiring suppliers to comply with specific quality standards – often accompanied by labelling and/or trade-marks – is often presented as an appropriate strategy to reach the desired goal. What remains unclear, however, is to what extent and on what grounds these standards can be considered sustainable. This paper presents findings of a study that aims to define what standards say they do and whether they do what they say.

In the 1990s the international trade in high-value agri-food products expanded significantly. Fresh and processed fruits and vegetables, fish, meat, nuts and spices, as a whole, now account for more than 50 per cent of developing country agricultural and food exports (Henson and Jaffee, 2006). Simultaneously, broader demographic and social trends have changed the nature of food markets in industrialised countries and altered expectations and demands (Buzby, 2003; Caswell, 2003; Henson and Jaffee, 2006). Consumers in industrialised countries have well-established and persistent concerns about certain dimensions of safety and quality attributes of food, especially with regard to environmental, social and food safety issues.

As a response to these concerns, governments and large food retailers, food manufacturers, and food service operators have developed and enhanced forms of sustainability controls, resulting in a complex and dynamic landscape of quality standards, trade marks and labels that continues to evolve (Busch, 2003; Ponte and Gibbon, 2005; Fold and Pritchard, 2005; Van Tulder with Van der Zwart, 2006). These public and private standards govern contemporary global agri-food systems and have become *a priori* mandatory, especially in high-value supply chains (Henson and Reardon, 2005). This has ambiguous effects on developing country suppliers: quality standards may either serve as a barrier or a catalyst to development (Henson and Jaffee, 2006; Danse and Vellema, 2007).

The public and private modes of supply chain governance have distinct institutional forms, but influence one another in terms of both content and procedures (Busch and Bain, 2004; Ponte and Gibbon, 2005; Ingenbleek and Meulenberg, 2006). It is therefore increasingly difficult to separate their distinct influence on firm conduct and supply chain sustainability performance, including the effects on suppliers. The web of partly overlapping quality standards puts both leading firms in the chain and NGOs in a dilemma when they have to choose or promote specific standards. These considerations led to the present study involving a comparison of quality standards relevant to European retail firms that source also in developing countries.

This paper examines the effectiveness of quality standards. It follows a particular approach in that it involves an examination of objectives of-, and likelihood of compliance with quality standards, rather than an assessment of effects achieved. The central question addressed in this paper is twofold:

1. How to measure the qualities of voluntary quality standards in terms of (a) their contribution to sustainability of global, agri-food supply chains, and (b) the likelihood of compliance?
2. What is the relative performance of the main voluntary quality standards in the retail, and which variables are influencing the qualities and performance of these standards?

The paper is organized as follows. First we discuss the governance of sustainability issues based on supply chain and business development theories (section 2), which leads to a number of propositions (section 3). In section four we explain the measurement instrument that was developed for examining voluntary quality standards. Section five presents the findings of the study involving 30 quality standards common to the European retail sector. Section 6 concludes and considers some managerial and policy consequences.

2. Governing sustainability within global retail chains

Quality standards are agreed criteria. They are "external points of reference by which a product or a service's performance, its technical and physical characteristics, and/or process and conditions under which it has been produced or delivered, can be assessed" (Nadvi and Wältring, 2004: 56). Standards consist of "rules of measurement" and a "system of classifications based on quantifiable attributes") (Jones and Hill, 1994). Quality standards may be mandatory, or voluntarily adopted. Mandatory standards are imposed by governments and intended to protect business or consumers within national jurisdictions from dishonest practices and risks. Voluntarily adopted, private quality standards intend to enhance inter-firm compatibility and may improve competitiveness of individual firms. Private standards serve the function of reducing transaction costs for buyers in the chain, particularly in situations where buyers have no ownership relation with suppliers, and/or when suppliers are located in countries with less stringent mandatory regulations. The ever more complex production relations resulting from global sourcing and increasing product differentiation make the use of quality standards inevitable to facilitate tight coordination and harmonization of norms and codification (Humphrey and Schmitz, 2004: 353; Navdi and Wältring, 2004: 53-56). However, technical efficiency and transaction cost reduction are not the only reasons for companies to voluntarily adopt quality standards, and this can best be explained in the context of the Global Value Chain (GVC) approach.

Like mainstream supply chain theories, GVC analysis uses the chain as metaphor for studying relationships between aligned firms. Its value added is the analytical model that (a) includes dependency relationships between small, local firms and multinational corporations that are linked together in global markets, (b) considers the influence exerted by non-market actors, such as governments and NGOs on the intra-chain relations, and (c) facilitates the analysis of normative issues related to production processes and the distribution of chain gains (see e.g. Gereffi and Korzeniewicz, 1994; Gibbon and Ponte, 2005; Humphrey and Schmitz, 2004).

Central in GVC analysis is the organizational or governance dimension, which refers to forms of intra-chain coordination. In most global value chains a 'lead' or 'core' company can be identified on the basis of its market power and value-added generation. The core company is increasingly strengthening its position as chain leader or chain director by setting the standards for quality of production for the entire chain. Quality management has in fact become one of the key aspects of chain governance (Fulponi, 2007; Gibbon and Ponte, 2005; Muradian and Pelupessy, 2005). It is a source of power over other companies in the chain which is not related to the degree of ownership, but rather to the ability of managing the sales of the chain's product. Designing quality and performance criteria is a form of hands-off coordination, for which the costs are often transferred to suppliers. Hence, the voluntary character of a private standard is limited to the core firm who deliberately chooses for its adoption. From the perspective of primary suppliers in the value chain, private quality

standards are *de facto* mandatory; they run the risk of losing business if they fail to comply with them.

In global food chains, the retail sector has assumed core positions with the prospect that, in 2010, only ten (food) retail firms are expected to dominate the supermarket sector worldwide (Fox and Vorley, 2004). The food chains also have become more international/global due to the expansion of imported food items. For example, 6 per cent of Wal-Mart's total merchandise was imported in 1995, against 60 per cent in 2004. Companies like Wal-Mart have become gate keepers to the main global markets.[2] The flip site is that, as the core firms in global chains, retail companies are being held accountable and responsible for the product qualities and the conditions under which they were produced wherever in the chain. To avoid the risk of reputation damage and extended liability, retail firms govern their supply chains intensively by means of a wide range of voluntarily adopted quality standards. These standards are intended to make supply chains more sustainable, which means that they consider the interests of people, planet and profit (Elkington, 1998; Serageldin, 1996; Serageldin and Steer, 1994). Most of the voluntary quality standards in the retail sector focus on socio-economic conditions of production, environmental protection or food safety. However, little is known about the extent to which these standards cover the sustainability categories and whether or not the requirements are met.

Key in the assessment of the qualities of quality standards is their compliance, that can be defines as "rule-consistent behaviour of those actors, to whom a rule is formally addressed and whose behaviour is targeted by the rule" (Börzel, 2000). Since it is almost unfeasible to verify compliance for all product parts at all stages in the value chain for a large number of quality standards, the effectiveness of quality standards could also be assessed by an examination of the *likelihood* of compliance as a proxy of real compliance.

'Compliance likelihood' is the probability that companies will conform in practice to their private codes or quality standards, and that the claims will in fact be translated into responsible behaviour and action (Kolk and van Tulder, 2005). The compliance likelihood is determined by the compliance mechanisms included in codes or quality standards and the extent to which the claims put forward are measurable. The more specific the codes or quality standards are, the better can they be measured and, subsequently, monitored. Monitoring is expected to enhance the standards' comprehensiveness and compliance likelihood.

One of the variables that are likely influencing compliance likelihood of voluntary quality standards is the *standard initiator*. A few publications have introduced typologies of standards based on the designers: 'standard initiators' (Muradian and Pelupessy, 2005), and standard 'key drivers' (Nadvi and Wältring, 2004). Both classifications serve as empirical categories. They group standards but they are not based on conceptual models intended to explain relationships with the standards' qualities. International standard initiators can be fivefold: (1) individual firms, (2) consortia of firms or branche organizations (so called Business-Support Groups, BSGs), (3) non-governmental or non-profit oriented organizations, (4) international governments, and (5) partnerships between firms, NGOs and governments. The concept of corporate social responsibility (CSR) approaches[3] (van Tulder with van der

[2] "Is Wal-Mart good for America?" Interview with G, Gereffi. *Frontline*, November 16, 2004, www.pbs.org/wgbh/pages/frontline/shows/walmart/interviews/gereffi.html

[3] CSR is the principle that companies are responsible for solving problems they have caused, and for helping to solve problems that are indirectly connected to their core functions (see e.g. Wartick and Wood 1998: 76).

Zwart, 2006; van Tulder *et al,* in press) employs a similar typology but is more promising in terms of proposition development. Four types of core firm CSR strategies can be distinguished in value chains, determined by two dimensions: firm responsiveness and the firm's moral attitude. An 'inactive' CSR approach is characterized by an inside-out (shareholder) orientation and a corporate focus on liability rather than moral responsibility. The opposing CSR approach is called 'pro-active', which has an interactive (stakeholder) orientation and is focused on both liability and moral responsibility, while engaging other stakeholders in the process of standard development. A 're-active' CSR strategy is an outside-in perspective and goes beyond liability towards responsibility, primarily because of potential reputational damage. A re-active CSR approach often goes together with relatively lax compliance rules. An 'active' CSR strategy first considers the moral responsibilities of firms and tries to engage other stakeholders in adopting its own standards. The standard might be relatively vague, but its compliance rules are often rather stringent (responsibility and liability are the same).

3. Propositions

There are four dimensions of the standardization process that can be taken into consideration when assessing the quality of the standards: (1) who initiated the standard and their position in the GVC; (2) whether this is done by an individual actor or a group of actors and what the impact of timing on the standard is; (3) the focus of the standard in general and (4) the issues in specific that are addressed by the standard. For each of these dimensions we have developed propositions.

First, as regards, the initiator of the standard, the approach taken can be expected to be derived from the relative position inside or outside the global value chain. Standards initiated by individual NGOs can be expected to show the highest possible compliance likelihood, compared to standards designed by firms themselves. NGOs in general will opt for greater transparency of the whole value chain, since they derive their bargaining position and legitimacy directly from citizens and consumers. In the literature on codes of conduct (cf. Kolk and Van Tulder, 2005) NGO codes have previously been found to have the highest compliance likelihood. NGOs, however, operate more or less outside of the GVC, which makes it easier to adopt an active CSR approach and operationalise standards with a higher compliance likelihood. Intergovernmental organizations, also derive their legitimacy from citizens – which would make their orientation comparable to that of NGOs - but in a more indirect manner, i.e. through their governments. This makes the compliance likelihood of their standards lower than for individual NGOs.

Individual food producers operate of course inside the value chain. They are generally positioned downstream of the retailers and upstream from the suppliers of resources. So, they have a mixed position in the GVC which also could lead to a more ambiguous attitude towards standards. They are on the one hand interested in managing their own value chain, which is enhanced by keeping much of the chain information proprietary. On the other hand, they are in direct contact with the retailers that ask for the biggest possible transparency. So their position in the value chain will probably lead to considerably lower compliance likelihood than NGO standards. However, as compared to (international) governments, it is not really clear what can be expected. The pressure on firms to (at least) adopt some of the minimum standards that are introduced by intergovernmental agencies is high, so it is unlikely that major firms – even if they adopt only a reactive approach - will go below the international

standard. Leading companies that for instance adopt quality standards from NGOs do so in the context of an active or pro-active CSR strategy. Their participation in the design likely also results in a standard that is more apt to the firm's organization and long-term strategy. Furthermore, since these standards have been co-designed by non-private stakeholders that are not primarily focused on the business case, they presumably entail more opportunities for sustainable supply chains than do business-developed standards.. .

Proposition 1: Quality standards designed by individual NGOs, have higher compliance likelihood than have standards designed by individual firms. Quality standards pioneered by individual firms will have at least the same compliance likelihood as international government standards.

Since the beginning of the 21st century, the standard formulation and implementation process by individual actors has become complemented by more partnership oriented processes (Van Tulder, v.d. Zwart, 2006). Partnerships with stakeholders are a necessary, but not sufficient condition for a pro-active approach. Partnerships have generally been struck within the sector (through industry associations), between NGOs and firms and by intergovernmental organisations, often in the form of multi-stakeholder dialogue processes. The partnership process can include consultation or joint standard setting. The latter requires much more commitment of the participating actors than the former. Consultation are often part of a more re-active approach to standards. Standards struck within the same group, such as industry associations, have been found to go for the 'lowest common denominator' as regards standards and codes (ibid). The reason for this being that in industry associations, the least interested party dominates the standard setting process. This process has also been referred to as 'adverse selection' (Lennox and Nash, 2003). Standards initiated by business sector groups can in general be expected to particularly reflect corporate concerns for liability prevention and reputation damage. BSG standard will show relative low compliance likelihood. Partnerships with outside parties (bilateral or tripartite partnerships; see Kolk *et al*., 2008) that lead to standards have the potential to be easier accepted than standards independently proposed by these 'outside' parties. They could evade adverse selection effects or the lowest common denominator mechanisms that BSGs suffer from. Many of the intergovernmental organisations have actively engaged in stakeholder consultation with firms and NGOs. The coordination of this consultation, has been in the hands of the intergovernmental organization. Joint standard setting – rather than consultation - requires that firms take the initiative for the partnership or actively engage in the formulation of the standard by the outside party. The active involvement of firms can make the standard more specific, and thus increase their compliance likelihood. Most of the actual partnerships, however, are yet in their infancy and therefore suffer from considerable start-up problems. Bi/trilateral partnerships initially will suffer from comparable problems as BSGs (such as adverse selection and lowest common denominator). The compliance likelihood of the standard coming from partnerships depends on the strategic intentions of in particular the firms that participate in the partnerships and their position in the value chain. The more parties are represented and the more diverse their interests are, the slower the partnership will lead to specific and high quality standards.

Proposition 2: Quality standards designed by (bi/trilateral) partnerships have potentially a higher compliance likelihood than standards designed by individual NGOs, firms or inter-governmental organizations. Partnership standards that have experienced a longer gestation period, or that represent very diverse interests might still show relatively low compliance likelihood.

A third dimension of quality standards is their focus. Quality standards may either be focused on a specific product category or have a generic focus. The criteria of product-specific standards are focused on a specific product and are likely more detailed and technical in nature. This makes such standards easier to codify by the standard's organization and hence better to understand by the standard-adopting firm compared to generic standards. The higher level of specificity facilitates compliance likelihood.

Proposition 3: Product-specific standards show higher compliance likelihood than do generic quality standards.

Finally, this paper distinguishes three issue categories of sustainability, involving socio-economic, environmental and food safety issues. Earlier research on '*implementation* likelihood'[4] was found to be more advanced for environmental than for social issues (Kolk, 2004). Quality standards specialists explain this finding as follows.[5] Compared to socio-economic standards, environmental and food standards can be codified at a more quantifiable level due to their technical nature. For these standards it is easier to set verifiable indicators, while the technical language enables engineers, both the standard developers as well as the standard adopters, to speak a common language. Furthermore, environmental standards have a longer history of development compared to socio-economic standards that emerged only during the last decade.

Proposition 4: Quality standards that are primarily focused on environmental en food safety issues have higher compliance likelihood than have quality standards that focus on socio-economic issues.

4. Methodology

Measurement instrument
The measurement instrument (Table 1) is derived from the one that was introduced by Kolk, van Tulder, and Welters (1999) and Van Tulder and Kolk (2001) for the analysis of the compliance likelihood of company Codes of Conduct. Their instrument classifies company codes along the dimensions of specificity and compliance. The underlying assumption of the instrument is that the likelihood of the firm's compliance to its own Code of Conduct is correlated with the specificity of the code's issues, the measurability of the plans, and the transparency of the code's monitoring system and compliance procedure. The instrument is also relevant for the analysis of quality standards, although the differences between company codes and industry quality standards require adjustments.

The specificity part of the measurement instrument comprises of the standards' scope, embeddedness and measurability. The sustainability *scope* is measured through 19 indicators of socio-economic, environmental, or food safety sustainability. Since standards may combine elements of these three sustainability categories, the instrument first measures how many of the 19 indicators are covered by the standard, resulting in a 'scope score'. The scope also

[4] 'Implementation likelihood' is a variation of compliance likelihood that refers to the probability of implementation of codes of conducts within the company's organization (Kolk, 2004).
[5] The following argument is based on private communications with Hans Blonk (Blonk Milieu Advies, the Netherlands) and Jan Willem van der Schans (Agricultural Economics Research Institute, the Netherlands) on 26 September 2007.

provides an indication of the main sustainability focus of the standard, i.e. the degree to which a standard concentrates on indicators of one sustainability category. Standards may only minimally mention specific sustainability issues because for these aspects they resort to third standards. For this reason the *embeddedness* of standards in third standards and the degree of such embeddedness are measured separately. The reference to third standards may only be general, i.e. without a clear indication of the way the standard relates to this third standard, or may include an explicit adoption of the third standard.[6] It is assumed that a general reference to, or the adoption of third standards enhances compliance likelihood, because (a) the specificity of the standard increases and (b) an additional compliance mechanism (of the third standard) is involved. With *measure* the standards can be checked for measurable specifications. The better standards can be measured, the better the opportunities for monitoring, and the higher the likelihood of compliance (Kolk *et al*., 1999, pp. 153-4).

The compliance part of the instrument is inspired by the 'managerial approach' rather than the 'enforcement approach' to compliance, which embraces a tough coercive strategy of monitoring and sanctions (Tallberg, 2002). The managerial approach advances compliance through "jawboning", i.e. pressure through strong persuasion, which involves problem-solving based on capacity building, rule interpretation and transparency (Chayes and Chayes, 1993). The instrument explicitly measures transparency and support, and values assistance as a sanction to non-compliance higher than immediate delisting. The *transparency* criterion measures the public availability and clarity of information on procedures and costs through the Internet and hence the accessibility of this information to potential new members and consumers. *Support* indicates the existence and intensity of (financial) support that is provided to the standard's member firms. The *monitoring* of standards involves four aspects that together cover the observance and examination or auditing of compliance by standard members. The manner in which the standard manages non-compliance is measured through four distinct aspects related to *sanctions*.

Except for scope, all criteria and indicators assess qualities of quality standards that are relevant to their compliance. The assumption underlying the instrument is that compliance is fostered by adoption of third standards, by measurable specification of targets, by procedural transparency, support of member firms, clear, frequent and low cost monitoring by independent organizations, and clear and reasonable sanction procedures. In other words: the higher the total score, the higher the likelihood of the compliance with the standard by member firms.

[6] Referencing to third standards may also have other reasons. The EurepGap (GlobalGap) website for instance refers to at least 12 third, national and international quality standards because they are accredited by EurepGap as standards that meet EurepGap criteria. Even tough such accreditation likely contributes to the international recognition of the EurepGap standard, these references enhance the compliance likelihood of the third standards, not of EurepGap itself. For this reason they were considered not to be relevant to this study

Table 1: Measurement framework for the analysis of quality standards

		Criteria		Indicators	Classification
SPECIFICITY	SCOPE	Socio-economic issues	1.Employment 2.Training 3.Working Conditions 4.Industrial relations 5.Force 6.Human rights 7.Business interest 8.Fair trade	Employment promotion; equality of opportunity and treatment; job security Workers Wages and benefits; condition of work and life; safety and health Freedom of association; collective bargaining; consultation; examination of grievances; procedures for dispute settlement Reduction/elimination child labour; forced labour; disciplinary practices Universal declaration human rights, including the protection of cultural values Reducing trade barriers, promoting trade; promoting technology; innovation; efficiency Fair payment; bribery; pricing; market requirements; transparency, cooperation	Ranging 1-8 / 8
		Environmental issues	1.Water management 2.Waste management 3.Chemical use / storage management 4.Energy management 5.Animal welfare 6.Nature conservation	Re-use of water, reduction of use Packaging materials; dumping; hazardous waste Controlled use of agro-chemicals Re-use of energy; green energy; reduction of usage Protection of animals, mistreatment, animal rights, painless production methods Land use and soil, natural production environment, biodiversity, no GMOs	Ranging 1-6 / 6
		Food Safety issues	1.Hygiene 2.Chemical residues 3.Product manipulation 4.Quality management 5.Traceability	Temperature; reducing bacteria; handling procedures (e.g. gloves), equipment and materials used Listed pesticides; reducing risk of residues; contamination; use of medication for animals Restrictions GMOs, nutritional value, additives Improving product, production, process Requirements for end products; display of labels, clearness, customer satisfaction	Ranging 1-5 / 5
			Standard's total scope score		**Total / 19**
	EMBEDDEDNESS	Reference to third standards	Standards may have enhanced their focus by embedding themselves for specific aspects in existing, international third standards Third standards may (1) be referred to in a general way (i.e. only mentioning the standard), or (2) be referred to through an explicit adoption of the entire standard		1 = general reference; 2 = explicit adoption third standard
	MEASURE	Specification of issue focus	To what extent does the standard provide measurable specifications? Such specifications may (0) not be available, (1) only be included in parts of the standard, (2) be included in most parts, or (3) be available for all parts of the standard.		0 = NA; 1 = in parts; 2 = most of standard; 3 = entirely
COMPLIANCE	TRANSPARENCY	Transparency of standard	Is detailed information about the standard (requirements, costs, compliance procedure, support) available *online* and accessible for interested persons external to the value chain (consumers e.g.)?		0 = no; 1 = in part; 2 = Yes
		Cost transparency	How transparent are the costs for certification? Transparency may range from (0) zero / no information, (1) vague, (2) clear to vague, to (3) clear.		0 = NA; 1 = vague; 2 = vague to clear; 3 = clear
	SUPPORT	Level	Do standard organizations support value chain actors in the implementation of the standard? Standard organizations may (0) not, or only provide basic guidelines, (1) offer passive educational materials without additional support service, or (2) provide for active and specific support (training, expertise)		0 = Basic; 1 = Education; 2 = active Support
		Financial support	Who pays for the support? There are two options: (0) value chain actors have to pay for themselves, or the standards fail to provide online information on this issue, and (1) the standard organization or a third party is covering (a part of) the cost for support		0 = chain actors or unclear 1 = standard org or 3rd party
	MONITORING	Process of monitoring	Is there a monitoring process specified? This process may be (0) unclear or not provided for, (1) only vaguely referred to, (2) clear to vague with only clarity in some parts of the process, or (3) clearly described.		0 = NA; 1 = vague ; 2 = vague to clear; 3 = clear
		Cost of certification	Who assumes the costs for the monitoring process? This may be (0) entirely unclear, (1) the member firm itself, or (2) a third party.		0 = unclear; 1 = firm; 2 = third party
		Position monitoring organization	How independent is the monitoring organization? The organization might be (0) not existing or unclear, (1) a member firm, (2) a business sector group (BSG), or (3) a governmental authority, the standard's organization, or an external organization		0 = unclear; 1 = firm; 2 = BSG; 3 = governmental/ standards'/ external organization
		Frequency monitoring	How often is compliance to the standard monitored? Monitoring may (0) be unclear, or take place (1) less than yearly, (2) yearly, or (3) more times per year		0 = unclear; 1 = less once/year; 2 = yearly; 3 = more once/year
	SANCTIONS	Sanctioning process	How detailed and complete is the description of the sanctioning process? Standards may (0) fail to provide information on a sanctioning process, or offer (1) a general, vague reference only, (2) insight in some parts of the process, or (3) a detailed and a complete description		0 = NA; 1 = general, vague; 2 = partial; 3 = detailed and complete
		Compliance flexibility	What happens if member firms fail to comply with the standard's rules? Standard may (0) be silent on this situation, or respond by (1) immediate delisting, (2) a warning first, or (3) providing assistance to further compliance a.s.a.p.		0 = NA; 1 = delisting; 2 = warning; 3 = assistance
		Publicity on sanctions	Does the standard organizations publicize the failure to meet the standards? Standard may (0) fail to provide information on this issue, (1) publicize this internally (e.g. only on own website, or only by updating members list), or (2) externally (e.g. by press statement or advertisements)		0 = NA; 1 = Internally; 2 = Externally
		Fraud protection	Does the standard have procedures in place to protect its members from misuse or fraud of the standard's label? Standard may (0) fail to provide information on fraud protection procedures, (1) refer to these procedures in a general way, or (2) provide a detailed insight in their procedures.		0 = NA; 1 = general; 2 = detailed

NA= Not available

Sample of standards
The study is based on a sample of 30 quality standards (Table 2). They include voluntarily adopted quality standards designed by either public or private entities, and they are selected because of their relevance to the food retail in Europe. The sample is relatively extensive as it is approaching the entire number of voluntary standards that European retailers apply to their suppliers located in developing countries.

Data collection
The data on individual standards are collected through the Internet only, by extensively searching the websites of the standards in the sample. Our data collection strategy was motivated by the assumption that analysis of qualities of quality standards could best take place by relying on key information that is publicly available online. At this stage of the study none of the 30 standard organizations was contacted for additional information, and no experts were consulted in respect of the objectives or procedures of specific standards.

5. Main findings

General results
The results of the study are displayed in Tables 2 and 3. Table 2 identifies the individual standards, their main characteristics, and their classification in terms of sustainability scope and compliance likelihood. Table 3 indicates the main findings per category of quality standards.

The average scope of the 30 quality standards is 7.6, which represents around 40 per cent of the possible sustainability criteria (19) that are used in this study. This makes the standards, on average, relatively narrow in scope. Utz Certified has the broadest scope (15) and the European Fair Trade Association the lowest (2). The compliance likelihood average for all standards is low: the average score is 14.3, which amounts to 42 per cent of the maximum possible score of 34. The standard with highest compliance likelihood is EKO (28); the lowest score is for the Sustainable Consumption Roundtable standard (1).

Differences by standard initiator
With regard to scope we observe that standards initiated by individual NGOs are broadest in scope. They include almost twice as many environmental and socio-economic criteria (13) than do standards initiated by others. Standards initiated by (inter) governmental organizations define relatively more food safety items than the standards initiated by others. Product-specific standards address more environmental and socio-economic items than generic standards, while the score for food safety is almost even.

Compliance likelihood is highest with standards initiated by individual NGOs (19.7), followed by standards initiated by business sector groups (BSGs) (15.8). Standards initiated by partnerships have the lowest compliance likelihood score (11.8), followed by standards initiated by (inter)governmental organizations (12.6). This result largely supports the first proposition that quality standards designed by individual NGOs show higher compliance likelihood than standards designed by individual firms and/or international governments. The expectation as expressed in the second proposition was not yet supported by the existing findings: quality standards developed in present partnerships do not have higher but lower compliance likelihood than standards designed by individual NGOs. The compliance likelihood score for partnership-initiated standards is generally lowest, with the exception of

the Common Code for the Coffee Community standard, which has a high compliance likelihood score (21), well beyond the average of NGO standards. The code knows already a relatively long gestation period (since 2002) and has strong commitment from the (German and Swiss) government, which makes the contributing stakeholders rather cohesive. The latter hints at initial support for the second part of proposition two in which the duration and the diversity of the partnership is considered important for its qualities. The number of observations for this study is too limited, however, to reach more solid conclusions.

Differences by product group
The scope of the 19 standards that were labelled as 'product-specific' is with a score of 8.1 slightly broader than that of generic standards (6.7). This is somewhat surprising in that intuitively one would expect product-specific standards to be narrow in scope. The most obvious explanation is that a product focus enables standard initiators to define more explicit requirements. This argument is supported by the relatively high (2.26) score on measurability for product-specific standards, compared to a low score (1) for generic standards.

In respect of compliance likelihood, product-specific standards do much better than generic standards. The respective scores are 16.7 and 10.1. This finding confirms proposition 3. Moreover, all eight standards that are classified as standards with high compliance likelihood are product specific. It is likely that these high scores correlate with standard initiators. All standards initiated by individual firms or NGOs are product specific, while BSG- and partnership-initiated standards are either generic or product-specific standards. The majority of the (inter)governmental standards can be categorized as generic.

Sustainability focus
The fourth proposition holds that those quality standards that are primarily focused on environmental en food safety issues have higher compliance likelihood than have quality standards focussing on socio-economic issues. It is not possible to test this proposition. The data on sustainability focus in Table 2 shows that the far majority of quality standards combine criteria for more than one sustainability category. Only seven standards focus on one category, while ten standards comprise of criteria from all three categories. The majority of standards show a stronger focus on socio-economic and/or environmental issues. Utz Certified is not only the broadest in scope, but also the most balanced: it scores a 5 for all three categories. The latter, however, implies also a somewhat lower score on socio-economic and environmental issues than the NGO codes.

Standard scope
With regard to the influence of the standard's scope on compliance likelihood we had no prior proposition, apart from the intuitive idea that, combined with high compliance likelihood, a broader set of sustainability requirements leads to a higher sustainability effect. On the other hand, again intuitively, one would expect broader standards to have lower compliance likelihood because a wide range of objectives are likely more difficult to achieve than few objectives. From the sample, 14 standards can be categorized as having a narrow scope (2-6.3), while the remaining 16 standards are equally divided over the categories medium scope (6.3-10.7) and broad scope (10.7-15). Table 3 shows that the standards with a broad scope have higher compliance likelihood than standards with a narrow or medium scope. However, Table 2 indicates that all standards with high compliance likelihood are either narrow or broad in scope. This ambiguous finding could indicate that a third, yet unknown variable influences the effect of the standard scope on the compliance likelihood.

Table 2 Classification of individual standards by initiator, product focus, sustainability focus, scope and compliance likelihood

Standard name	Initiator	Product focus	Socio-economic(8)	Environmental (6)	Safety (5)	Scope	Compliance likelihood
Nature and More	Individual firm	Product specific	5	2	3	Medium	Low
Sara Lee's Supplier Selection Guidelines	Individual firm	Product specific	5	2	0	Medium	Low
Starbucks Green Coffee Purchasing Guidelines	Individual firm	Product specific	3	5	0	Medium	Medium
Marine Stewardship Council	Individual firm	Product specific	0	2	0	Narrow	High
Utz Certified	Individual firm	Product specific	5	5	5	Broad	High
European Fair Trade Association	BSG	Generic	2	0	0	Narrow	Low
Ethical Tea Partnership	BSG	Product specific	5	0	0	Narrow	Medium
Global Food Safety Initiative	BSG	Product specific	0	0	3	Narrow	Medium
EUREPGAP Code (1)	BSG	Product specific	0	3	2	Narrow	High
British Retail Consortium code	BSG	Generic	6	3	3	Broad	Medium
QS	BSG	Product specific	0	4	2	Narrow	High
Sustainable Agriculture Initiative Platform	BSG	Generic	4	5	2	Broad	Low
MPS-A,B,C & SQ (2)	BSG	Product specific	5	4	0	Medium	High
Good Manufacturing Practices	(inter)governmental	Generic	0	0	3	Narrow	Medium
EKO	(inter)governmental	Product specific	0	0	2	Narrow	High
FAO Code for responsible fisheries	(inter)governmental	Product specific	3	3	0	Medium	Low
OECD Guidelines for Multinational Enterprises	(inter)governmental	Generic	7	2	0	Medium	Medium
Global Compact	(inter)governmental	Generic	7	4	0	Broad	Medium
ISO 22000	(inter)governmental	Generic	0	0	3	Narrow	Medium
ILO Convention 184	(inter)governmental	Generic	3	3	0	Narrow	Medium
FAO Code of Conduct on Pesticides	(inter)governmental	Generic	0	4	3	Medium	Low
Codex Alimentarius	(inter)governmental	Product specific	0	0	5	Narrow	Low
Rain Forest Alliance/SAN	Individual NGO	Product specific	6	5	0	Broad	Medium
Max Havelaar	Individual NGO	Product specific	7	6	0	Broad	High
Fairtrade Labelling Organisation	Individual NGO	Product specific	7	6	0	Broad	High
Sustainable Consumption Roundtable	Partnership	Generic	0	2	0	Narrow	Low
Ethical Trading Initiative Base Code	Partnership	Generic	5	0	0	Narrow	Medium
Fair Flowers & Plants	Partnership	Product specific	0	4	0	Narrow	Medium
Common Codes for the Coffee Community	Partnership	Product specific	7	5	0	Broad	High
Forest Stewardship Council	Partnership	Product specific	6	4	0	Medium	Medium

Table 3 Sustainability and compliance likelihood of 30 voluntary quality standards for the European retail

	Sustainability scope (Maximum score 19)	**Compliance likelihood** (Maximum score 34)
Overall mean	7,6	14,3 (42%)
By initiator		
Firm (N=5)	8,	14,4 (42%)
BSG (N=8)	6,8	15,8 (46%)
Government (N=9)	6,3	12,6 (37%)
NGO (N=3)	13,0	19,7 (58%)
Partnership (N=5)	7,0	11,8 (35%)
By product focus		
Product specific (N=19)	8,1	16,7 (49%)
Generic (N=11)	6,7	10,1 (30%)
By scope		
Narrow (N=14)		14,9 (44%)
Medium N=8)		10,1 (30%)
Broad (=8)		17,5 (51%)

6. Discussion and conclusion

The aim of this study was to determine the effectiveness of voluntary quality standards in contributing to the sustainability of global agri-food supply chains. We first developed a measurement instrument to assess the 'qualities' of quality standards required by the European retail sector in respect of their food suppliers from developing countries. Based on specificity and compliance indicators we could analyze and compare the standards on their sustainability scope and compliance likelihood. The results indicate some relationships between standard initiator, scope, product focus on the one hand, and sustainability scope and compliance likelihood on the other.

Average compliance likelihood for the entire sample of standards is - with a score of 42 per cent - quite low. So, the effectiveness of good intentions laid down in voluntary quality standards may be far lower than is often expected. But there is substantial variation among standards. Around 30 per cent of the standards in the sample show relatively high compliance likelihood. They include: Max Havelaar, Fairtrade Labelling Organisation, Utz Certified, Common Codes for the Coffee Community, EUREPGAP, QS, EKO, Marine Stewardship Council, and MPS A,B,C, & SQ[7]. These nine are the only standards in the sample that have a compliance likelihood score higher than 60 per cent. On average, their score is 70 per cent, with EKO on top (82 per cent).

[7] In this study MPS ABC and MPS SQ are considered to represent a single standard because publicly accessible sources create the impression that the requirements are part of one and the same standard. Since MPS ABC and MPS SQ have a distinct focus on respectively environmental and socio-economic issues, our decision may cause bias in the analysis.

Standards initiated by NGOs show highest compliance likelihood, followed by BSG-initiated standards. This positive result for NGO standards may be partially biased due to the small number of NGO standards in the sample (3), for which further research is recommended. BSG-initiated standards result in higher compliance likelihood than those initiated by individual firms or multi-stakeholder partnerships. This result may be explained by the fact that both the firms and partnerships represent two extreme sides of cooperation, but also by the relatively recent nature of many of the partnership initiatives. Individual firms might lack a clear view of sustainability and compliance requirements which a standard must include to improve its compliance likelihood. It is also possible that firms communicate their standard's requirements through non-public communication channels to their suppliers. Another explanation is that, despite the intentions that are communicated, standards designed by individual firms serve other means than improving the sustainability of the supply chain.

The reason that partnership standards result in low compliance likelihood may be that multi-stakeholder agreements can only be reached through consensus and compromise. This may diminish the vigour of the standard, an outcome that is reflected in a lack of detailed requirements on sustainability specificity and compliance. Partnerships with stakeholders outside of the direct value chain are generally considered to be a prerequisite for a more pro-active CSR strategy which potentially could render a much higher compliance likelihood, and thus sustainability, than existing (individual) initiative. From the results so far we can conclude that there is still a 'world to be gained' in the effective management of partnerships. However, the Common Codes for the Coffee community shows an interesting leading example which has already resulted in considerably higher quality levels than any of the other standards.

The interesting question now is whether the study allows us to make judgements on whether some standards are better than others. We think it is possible. When "better" is defined in terms of two qualities, a 'broad' sustainability scope (score >10.7) combined with relatively 'high' compliance likelihood (score >19), there are four standards that fit this definition: Max Havelaar, Fairtrade Labelling Organisation, Utz Certified, Common Codes for the Coffee Community. These are the better voluntary quality standards of the European retail; they have the most wide sustainability scope *and* the highest likelihood that they do what they say they do. So, the managerial conclusion of the above still is that either NGO-initiated standards or well managed partnership standards might provide the best guarantee for better and more sustainable standards that do not serve specific interests in improving the bargaining position in the value chain, but might lead to increased sustainability of the whole (global) value chain.

All these conclusions must be related to the limitations of this study, however. All data used were derived from publicly available, online sources, and the study focused on intentions and the probability of compliance, rather than the actual effectiveness of standards. To further validate the findings of this study, further research is therefore required that includes primary information sources other than those available online, expert opinions, and interviews with the actors in the global retail value chains actors.

Finally, our sustainability focus was based on the standards' requirements related to the three P's. We observed that the standards barely refer to requirements that facilitate the firm applying to develop profit generating activities in order to cover the additional costs of sustainability while at the same time remaining competitive. These profit generating activities could include the creation of value, competition reduction, and pricing. Literature on the impact of quality standards on suppliers' upgrading opportunities is extremely scarce (cf. Humphrey and Schmitz, 2004), and the area is largely under-researched. We realize that also this study on sustainability scope and compliance likelihood did not yet address these important issues. Further research should therefore aim at the relation between the qualities of quality standards and their effects on adding value within the GVC connecting food suppliers from developing countries to the European (food) retail sector.

References

Börzel, Tanja A, 2000. Private Actors on the Rise? The Role of Non-State Actors in Compliance with International Institutions. Max Planck Institute, Bonn. MPI Collective Goods Preprint No. 2000/14. Available at SSRN: http://ssrn.com/abstract=267733

Busch, L., 2003.The Moral Economy of Grades and Standards, J. of Rur. Stud.,16, 273–283.

Busch, L. and Bain, C., 2004. New! Improved? The Transformation of the Global Agri-Food System, Rur. Soc., 69 (3), 321–346.

Buzby, J. (ed.), 2003. International Trade and Food Safety: Economic Theory and Case Studies, Washington DC: United States Department of Agriculture.

Caswell, J., 2003. Trends in Food Safety Standards and Regulation: Implications for Developing Countries, in: L. Unnevehr (editor), Food Safety in Food Security and Food Trade, Washington, DC: International Food Policy Research Institute.

Chayes, A., and A.H. Chayes, 1993. On Compliance. International Organization, 47 (2), 175-205

Danse, M.G. and S. Vellema, 2007, "Small scale farmer access: to international agro-food chains: a BoP based reflection on the need of social embedded innovation in the coffee and flower sector". Greener Management International, Issue 51, 39-52.

Elkington, J, 1998, Cannibals with Forks: the triple bottom line of the 21th century business, British Colombia, Canada: New Society Publishers.

Gereffi, Gary and Miquel Korzeniewicz (eds), 1994. Commodity Chains and Global Capitalism. Praeger Publishers, Westport, CT, USA.

Gibbon, Peter and Stefano Ponte, 2005. Trading Down. Africa, Value Chains, and the Global Economy. Temple University Press, Philadelphia.

Henson, S. and S. Jaffee, 2006, Food Safety Standards and Trade: enhancing competitiveness and avoiding exclusion of developing countries, The Eur. J. of Dev. Res.,18 (4), 593-621.

Henson, S.J. and Reardon, T., 2005. Private Agri-food Standards: Implications for Food Policy and the Agri-food System, Food Pol., 30 (3), 241–253.

Humphrey, John and Hubert Schmitz, 2004. Chain governance and upgrading: tacking stock. In: Schmitz, Hubert (editors) Local Enterprises in the Global Economy. Issues of Governance and Upgrading. Edward Elgar, Cheltenham UK and Northampton, MA, USA, pp. 349-381.

Ingenbleek, P. and M.T.G. Meulenberg, 2006. The Battle between "good"and "better": a strategic marketing perspective on codes of conduct for sustainable agriculture, Agribus., 22 (4), 451-473.

Jones, E. and Hill, L.D., 1994. Re-engineering marketing policies in food and agriculture: issues and alternatives for grain grading policies. In: Padberg, D.I. (editor), Re-engineering marketing policies for food and agriculture. Texas A&M, 119-129. Food and Agricultural Marketing Consortium, FAMC, no. 94-1.

Kolk, A., 2004. More than words? An analysis of sustainability reports. New Academy Review 3, 59–75.

Kolk, A., R. van Tulder and C. Welters, 1999. International codes of conduct and corporate social responsibility: can transnational corporations regulate themselves?, Transnat. Corp., 8(1),143-180.

Kolk, A and R. van Tulder, 2002, Child Labour and Multinational Conduct: A Comparison of International Business and Stakeholder codes, J. of Bus. Eth. 36, 291-301.

Kolk, A and R. van Tulder, 2005, Setting New Global Rules? TNCs and Codes of Conduct, Transnat. Corp., 14(3),1-17.

Kolk, A., R. van Tulder and E. Kostwinder, 2008, 'Partnerships for development', European Management Journal, forthcoming

Kooiman, J., 1993. Modern Governance: New Government-Society Interactions, London: Sage.

Lenox, J.M. and Nash, J., 2003. Industry self-regulation and adverse selection: a comparison across four trade association programs. Business Strategy and the Environment 12 (6), 343-356.

Muradian, Roldan and Wim Pelupessy, 2005. Governing the coffee chain: the role of voluntary regulatory systems. World Dev., 33 (12), 2029-2044.

Navdi, K. and Wältring, F, 2004. Making sense of global standards. In: H. Schmitz (editor), Local Enterprises in the Global Economy: Issues of Governance and Upgrading, Edward Elgar, Cheltenham.

Ponte, S. and Gibbon, P., 2005, Quality Standards, Conventions and the Governance of Global Value Chains, Eco. and Soc., 37 (1), 1–31.

Serageldin, I., 1996, Sustainability and the world nations, first steps in an ongoing journey. Dev. Stud. and Mon. Series No. 5, World Bank, Washington.

Serageldin, I. and A. Steer, 1994. Making development sustainable: from concepts to action, World Bank, Washington, DC.

Tallberg, J., 2002. Paths to Compliance: Enforcement, Management, and the European Union. International Organization, 56 (3), 609-643.

Van Tulder, R. and A. Kolk, 2001. Multinationality and Corporate Ethics: Codes of Conduct in the Sporting Goods Industry. J. of Int. Bus. Stud., 32(2), 267-283.

Van Tulder, R. with A. van der Zwart, 2006, International Business-Society Management: Linking Corporate Responsibility and Globalisation, Routledge, London.

Van Tulder, Rob, Jeroen van Wijk, and Ans Kolk, in press, From chain liability to chain responsibility: MNE approaches to implement safety and health codes in international supply chains, J. of Bus. Eth.
Wartick, S.L. and Wood, D.J., 1998, International Business & Society. Blackwell, Oxford, UK.