
Improving sustainability in global supply chains with private certification standards: testing an approach for assessing their performance and impact potential

Walter J.V. Vermeulen* and
Janneke A. Metselaar

Copernicus Institute of Sustainable Development,
Utrecht University,
Heidelberglaan 2, 3584 CS Utrecht, The Netherlands
Email: w.j.v.vermeulen@uu.nl
Email: janneke.metselaar@gmail.com

*Correspondence author

Abstract: Sustainable supply chain governance approaches aim for improvement of environmental and community living conditions at the developing country's side of the global supply chains. Impact evaluation in remote and multiple sourcing countries is hardly done in practice because of its complexity and costs. Evaluating the impact *potential* with this methodology enables an easier route for assessing and comparing the fast growing number of standard systems. This article describes a methodology for assessing the *system performance* and *impact potential* of 'sustainable supply chain governance (SSCG) systems' and uses two cases of certifying sustainable products in the coffee sector (Fairtrade and Utz Certified) for verification of the method. The approach rests on the assumption that a certification system can have substantial impact reduction if it is well organised in its content and organisation: the governance system *impact potential* (GSIP). Two exemplary pilot cases show that changes made over the years in both systems are directly visible with the methodology and it can be used for monitoring the progress in markets for sustainable products.

Keywords: sustainable development; coffee; governance; supply chain; Fairtrade; Utz Certified; developing countries.

Reference to this paper should be made as follows: Vermeulen, W.J.V. and Metselaar, J.A. (2015) 'Improving sustainability in global supply chains with private certification standards: testing an approach for assessing their performance and impact potential', *Int. J. Business and Globalisation*, Vol. 14, No. 2, pp.226–250.

Biographical notes: Walter J.V. Vermeulen is an Associate Professor in Environmental Social Science at Utrecht University since 1996. His work focuses on design, implementation and effectiveness of new strategies for improving the environmental and socio-ethical performance of production and consumption systems by means of co-production of sustainable development strategies. He is also the President of the International Sustainable Development Research Society (<http://www.isdrs.org>).

Janneke A. Metselaar graduated in Sustainable Development: Environmental Policy and Management at Utrecht University. The research for her Master thesis focused on sustainable supply chains, for which she was positioned at the Netherlands Environmental Assessment Agency (PBL) in Bilthoven, Netherlands.

1 Introduction

With the processes of globalisation, the world economy has increasingly become international in scope and organisation, with firms linking together in worldwide ‘commodity chains’ or ‘supply chains’. In these supply chains the leading firms (buyers) are predominantly located in developed countries while production (suppliers) often takes place in developing countries in an unsustainable manner. To improve the conditions of production, both pioneers and leading firms have developed ‘sustainable supply chain governance (SSCG)’ systems, introducing product supply assurance standards, such as Fairtrade or Utz Certified. We define sustainable supply chain governance systems (in short: SSCG-systems) as “forms of cooperation of market actors in (international) supply chains (possibly together with non-market actors) in improving the environmental and social conditions of production operations in developing countries” (following Vermeulen and Seuring, 2009; Vermeulen, 2010).

Responding to increasing consumer demand and discourse in the public media on these issues, ‘bioneers and ecopreneurs’ (Schaltegger, 2002) in the market and civil society (NGOs) have been filling a ‘regulation vacuum’ with these approaches. Western governments cannot do much to prevent the increasing shift of environmental impacts towards developing countries, as it is the result of growing international trade. National Western governments are not entitled to directly address production conditions in developing countries. They have to walk the long route via supranational institutions (such as WTO, UN or OECD), with their weak implementation powers, and await the effective implementation of UN agreements by national governments.

We see increasing activity on these issues both in the market and in government policies. Various types of business-to-business SSCG systems have emerged in the last decade (Svensson, 2007; Seuring and Müller, 2008; Seuring et al., 2008; Vermeulen, 2010; Vermeulen and Kok, 2012). The *history* of global SSCG started with small ‘enlightened’ entrepreneurs, often with roots in civil society, who started by bypassing dominant mainstream value chains. Starting in the late 1970s and 1980s, the fairtrade initiatives began to create new and shorter value chains linking small producers in developing countries more directly with western consumers.

Examples of such first generation *single firm approaches* are the early NGO-based fairtrade firms applying certification (e.g., Solidaridad, which created the Dutch Max Havelaar brand), or firms applying procurement demands, as done by Xerox (McIntyre et al., 1998), Nestlé (Alvarez et al., 2010) and others (Vermeulen and Ras, 2006; Keating et al., 2008). These strategies grew into a second generation of *joint product chain arrangements*, where groups of businesses, civil society organisations or sector organisations cooperate in developing and applying sustainability standards for a specific product group (FSC for timber, Utz Certified for coffee, Rainforest Alliance for tropical timber).

A more recent development is a third strategy of *cross-sector joint product chain arrangements* where systems are developed covering a wide range of products (such as GlobalGAP). This allows market actors to reduce transaction costs for achieving sustainable production in the global trade context (Williamson, 2008) (see for a more detailed discussion Vermeulen, 2010).

The impact of these systems at the supply side is principally determined by the market share of these systems and their performance. Due to their complex nature, it is difficult to analyse and determine the final environmental and socio-economic impact at the supply side, which is often located in developing countries. In recent studies this was identified as one of the major gaps in the practice and analysis of self-governance in supply chains [Vermeulen et al., (2010), pp.59–61; Vermeulen, 2013]. A full system approach addressing the impacts in the whole product group is missing, allowing to sum the impacts of the various competing approaches active in one specific market, like discussed above. It would require extensive research in many supplying countries, which is hardly possible. But a more feasible route is possible. To enable an indirect assessment of this impact we propose a more feasible approach in Section 3. This approach rests on the hypothesis that a system can have a *potential* impact: the governance system impact potential (GSIP), based on the way it is organised and its contents.

In this line of reasoning the GSIP is the product of the market share and the governance system performance (GSP). The latter consists of three components determining the performance of the system (coverage of the sustainability aspects, precision of the standard and its compliance control). It is therefore assumed that both the system performance and a larger market share of a SSCG system (i.e., more sustainable products sold), will result in a larger potential impact of a system and that it may vary over time.

This brings us to the research question: “How can we describe the *potential* impacts of ‘sustainable supply chain governance (SSCG)’ systems, based on its core characteristics, rather than the measurement of these remote and diffuse impacts at the supply side?”

This article explains and justifies the methodology for measuring this impact potential and will illustrate its application with two pilot examples from the coffee chain¹.

In the next section we will first illustrate the development in supply chain governance in the coffee chain. In Section 3 we will explain the created methodology for analysing the GSIP. Then construction of the method and the different indicators in the method will be explained and a description is given on how the method should be used. To illustrate and verify the method, Section 4 presents two different case studies: Fairtrade and Utz Certified coffee. In Section 5 the conclusions of this article is presented.

2 The coffee supply chain

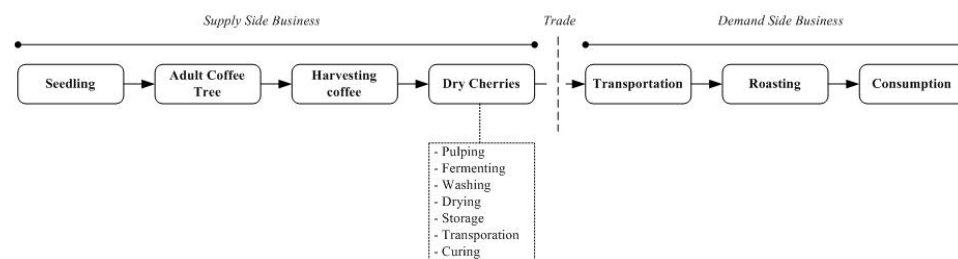
We have chosen the coffee supply chain as a case study to test the methodology, as coffee is a supply chain that is mainly controlled by the consuming countries in the North. Also, for coffee some of the oldest certification systems are available. While the consumption in these countries has increased, producing countries are facing a coffee crisis: there is too much production while farmers often receive too little money to cover

their production (Daviron and Ponte, 2005), implying that producing countries are not gaining much from the trading. This can lead to several socio-economic problems such as increasing poverty or loss of employment (Consumers International and IIED, 2005).

Besides these socio-economic problems, ecological problems also occur in coffee production. For instance, farms may be using different chemicals and pesticides (IISD, 2004; Consumers International and IIED, 2005; Gillison et al., 2004), which poses a threat to the natural environment, soil, surface water and wildlife.

The supply chain of coffee involves several phases (Figure 1): as coffee is grown from coffee trees, it takes about three to five years after sowing before the cherries growing on the tree can be harvested. After harvesting, the cherries are dried. This involves a process involving hulling, drying and peeling. After these processes the coffee is transported and traded to the import dock in the consuming country and further processed (roasted, packed) and consumed.

Figure 1 The coffee supply chain



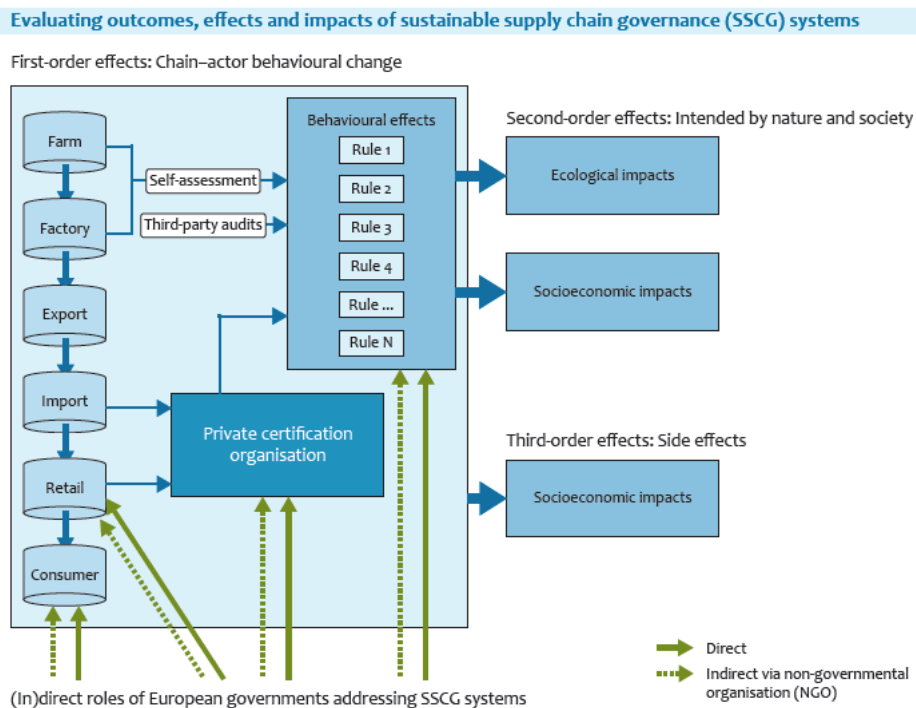
As result of the unsustainable situation in the coffee industry, sustainability has become a hot topic in the coffee industry. Different measures have been taken by firms to incorporate aspects of economic viability for the farmers, environmental conservation and increased social responsibility (Daviron and Ponte, 2005). One way of achieving more sustainability in the supply chain is by creating a certification standard (with labelling) such as, e.g., Fairtrade, EKO, Rainforest Alliance or Utz Certified. Sustainable *production*, as stated in Giovanucci and Koekoek (2003) ‘meets long term environmental and social goals while being able to compete effectively with other market participants and achieve prices that cover production costs and allows producers to earn an acceptable business margin’. Fairtrade coffee was first introduced in the Netherlands with Max Havelaar coffee in 1988. Since then, an increasing number of initiatives (code of conducts) started to certify coffee. Starbucks was the first to introduce its framework for a code of conduct in 1995. In that year also Rainforest Alliance started with its certification programme for coffee. An overview of these initiatives (codes of conduct) relevant for the coffee sector since Starbucks is provided by Kolk (2005). Her overview shows that the initiatives are different by nature: they are either multi-stakeholder initiatives, multinational corporation initiatives or non-governmental organisation initiatives. The mid 1990s showed more NGO action against flagship enterprises such as Philip Morris, Sara Lee, Nestlé and Procter and Gamble (Alvarez et al., 2010).

In a recent report we described the developments in the market for sustainable coffee in the Netherlands. Max Havelaar and Utz Certified are the most dominant private

standards. The sales of various forms of ‘sustainable coffee’ have long been at a low level of approximately 2–3%. The market shares of both organic and Fairtrade coffee on the Dutch market have remained stable over the recent years (Vermeulen et al., 2010; Vermeulen and Kok, 2012).

After 2003 more systems, linked to mainstream firms, have entered the market and this has resulted in a rapid growth of certified sustainable coffee (Coffee Coalition, 2007). In 2008, about 25% of coffee in the Dutch market was estimated to be certified (Fairtrade, Organic, Rainforest Alliance, Utz Certified, 4C, C.A.F.E., and AAA) (TCC, 2009). But such figures only show a part of the story: not all coffee from certified sources ends up as certified coffee on the shelves. Data on this justifies an estimation of an additional market share for unlabelled coffee sources from various certified growers (TCC, 2009), implying that markets are ahead of consumer demand for certified products. In addition to the 25–29% certified coffee another 15–20% coffee from certified grower is being sold at the market, summing up to approximately 45% in 2008 (Vermeulen et al., 2010). These developments make the coffee chain and its main private standards a good case for verifying the proposed methodology.

Figure 2 Evaluating outcomes, effects and impacts of SCG systems (see online version for colours)



Evaluating outcomes, effects and impacts of sustainable supply chain governance (SSCG) systems

Source: Vermeulen et al. (2010)

3 Methodology – governance system (impact) potential

3.1 Governance system impact potential

The certification systems as described in the previous paragraph are examples of SSCG systems. As discussed in Section 1, various types of business-to-business SSCG systems have emerged in the last decade. These SSCG systems claim to assure sustainable production practices at the remote supply side. Yet, the question always remains what factual final impact of these systems is on the ecology and community in the producing countries. Figure 2 illustrates the intended functioning of such governance systems, mainly aiming to establish the second order effects in nature and society.

Impact measurement takes place on the level of individual businesses in supply chains, either focussing on economic performance (Chan and Qi, 2003; Iles, 2007; Gopal and Cline, 2007; Kovács, 2008; Lozano and Huisingh, 2011; Youn et al., 2013), with a focus on certain aspects of environmental performance (Green et al., 1998; McIntyre et al., 1998; Tsoufas and Pappis, 2008; Schliephake et al., 2009; Testa and Iraldo, 2010) or on socio-ethical performance (Ciliberti et al., 2008). Cases of integrated assessment of both the environmental, social and economic aspects of sustainability are available (Keating et al., 2008), but again only on the individual businesses level. Still, the final impact (second order effects) is very difficult to determine due to various reasons. First of all: in most of these supply chains products are sourced from many different countries and regions, in many cases also difficult to access. Analysing the impacts would involve many different researchers and adjustment to local circumstances. Second: the aimed target, improving sustainability, in itself is very complex, including many different aspects, with different needs in types of data required. Third, adding to the complexity of determining the *exact* impacts of competing SSCG-systems is the issue that impacts can be caused by other developments as well (government interventions, autonomous market developments, technological change, etc.) which calls for filtering out such alternative explanations to establish cause-effect relations. Fourth, at the demand side there is competition and we see many diverse self-regulating initiatives. In this field the market is in the lead and with success, like we showed in the previous section. But the side effect of this is that there is not one single actor who is responsible for the whole picture. Looking at the international scientific literature, we indeed see a lack of data and only a rather small number of local case study-based articles. In most cases they focus on one system in one specific situation, like for example addressing only organic certification in one or two countries, like Albersmeier et al. (2009) did in Costa Rica and Brazil. Another example is an analysis of sustainability initiatives in the coffee supply chain and its effects on the plantations and surrounding communities by the International Institute for Sustainable Development (IISD). This study shows the various specific *actual* impacts of specific certification system at specific places and provides various well-grounded conclusions on how sustainability initiatives could improve, but is also not able to provide the full picture (Giovannucci and Purcell, 2008).

Even though these case studies are very useful, they hardly enable conclusions on the level of the full product chain. A second best option therefore would be to measure the *potential* impact on the ecology and community, which we call the GSIP. The GSIP expresses the probable relative share of a specific SSCG-system in the total intended final

environmental and socio-economic impacts at the supply side communities. It is thereby a proxy of what is assumed to be the impact in the supply chain. This impact potential is equal to the product of the GSP and the market share:

$$GSIP = \text{market share} \times GSP$$

Following a more detailed discussion by Vermeulen (2010), we can argue that the *performance* of SSCG systems depends on a limited number of characteristics of these governance systems. These include the following: ‘coverage and precision’, the ‘goal setting performance’, the ‘compliance control’ of the system, the provided ‘support’ for implementation by producers in developing countries, ‘knowledge transfer’ within the chain, the ‘openness and inclusiveness’ of the system and the ‘complexity of transactions’. Some of these characteristics may be seen as conditional for some of the other characteristics (like the goal setting determining the scope of the requirements) or as conditional for the final level of adoption by suppliers (like support and knowledge transfer being essential for the level of adoption). So, not all of these characteristics are needed to determine the performance of the governance system (as defined in Section 1). Therefore this methodology will focus on three core components, which we can assume to have the most direct influence upon the final impacts of the system in developing countries and not on the level of adoption by suppliers (because this will be measured by using market share data). In this way we use ‘coverage’, ‘precision’ and the ‘compliance control’ as the key features of a SSCG system to determine the likelihood of one single SSCG system to successfully improve sustainability at the supply side. Following this line of reasoning, the GSP can be determined by multiplying these three main performance components; compliance (C), precision (P) and compliance control (CC), thus:

$$GSP = C \times P \times CC.$$

To make this possible the three performance components need to be determined. A broader definition and explanation of these three components is provided in the following sections.

3.2 Coverage

Coverage addresses two crucial issues: does a SSCG-system address *all links* in the chain and for *all aspects* of sustainability? We can describe coverage as the degree to which all links of the supply chain are covered in terms of all aspects sustainability. Here it is essential to include aspects related to all three spheres of sustainable development (social sustainability, economic sustainability and environmental sustainability)². The indicators chosen to define coverage need to be applicable on a firm level as this is where production in a SSCG system takes place. As this study uses coffee as a case study, the indicators had to be applicable to coffee production as well, though these indicators could be used for analysing agricultural supply chains as well.

In order to identify the different indicators for determining the coverage, different theories on sustainability were applied to select appropriate indicators as well as literature about indicators for sustainable development (United Nations Statistical Division, 1997; United Nations Department for Sustainable Development, 2007). Indicators for the social and socio-economic sphere are based on the Daly triangle. The hierarchical pyramid of

the Daly triangle displays different ends and means at society level, connecting this to different kinds of capitals: natural capital at the bottom, via built capital and human capital to social capital and with 'well-being' at the top. We could translate these layers of the Daly triangle to firm level indicators suitable for the various indicator groups we included under 'scope' (Daly, 1990).

Table 1 Sustainability spheres, indicator groups and their indicators

<i>Indicator group</i>	<i>Indicator*</i>
<i>Social sustainability sphere</i>	
Poverty reduction	Housing availability Human rights Fair income
Health	Sanitation (availability) Drink water (availability) Healthcare (access)
Education	Schooling (provision)
Safety	Employee safety Product quality
Governance	Employee (discrimination) Reciprocity (no) Corruption
<i>Environmental sustainability sphere</i>	
Atmosphere	CO2 emissions Other greenhouse gas emissions Air pollutants (other) Ozone layer depleting substances
Terrestrial	Chemical use Agrochemical use Fertilisers Deforestation Soil protection
Water	Water quality Water quantity
Biodiversity	Species (preservation) Habitat (preservation) Preservation management Impact prevention
Waste	Waste generation Waste management and recycling
Resources	Energy use Energy availability (infrastructure)

Note: *For a specification on the selection of these indicators (see Metselaar, 2010)

Table 1 Sustainability spheres, indicator groups and their indicators (continued)

<i>Indicator group</i>	<i>Indicator*</i>
<i>Economic sustainability sphere</i>	
Employment	Community employment
	Female employment
	Child labour prevention
Labour conditions	Freedom of association (speech)
	Work load (per day)
Social security	Retirement arrangements
	Sick leave (insurance)
	Maternity leave
Income and finances	Credit provisions
	Fair transaction arrangements

Note: *For a specification on the selection of these indicators (see Metselaar, 2010)

The Daly triangle also provides a starting point for determining indicator groups for the environmental sphere, as the bottom of the pyramid represents the natural capital. The natural capital theory provides a basis for creating a set of indicators covering both issues of stock/capital and flow/income, as well as indicators that are part of a larger non-human system. Indicator groups were found by using both the bottom layer of the Daly triangle, and also the natural capital theory, as described by Costanza and Daly (1992). These indicator groups were also translated to firm level indicators.

In the end, 15 indicator groups were distinguished (Table 1) with 39 indicators². The standards (in this case, the standards of Utz Certified and Fairtrade) are analysed by applying these indicators. For the presence of each indicator, a score of 1 is given. A more detailed explanation of calculation is provided in Section 3.5.

3.3 Precision

Precision refers to the level of detail of prescriptions and to the requirement level in the different standards. While the ‘coverage’ concept describes the presence of the various sustainability aspects, the ‘precision’ concept expresses the way the different requirements are formulated: is it done in a precise and stringent way?

3.4 Compliance control

The third performance component measured is the compliance control. Here we can argue, (based on literature on regulatory enforcement) that the chance of a specific rule to be complied to will depend on the chance of inspection and the severity of the sanction. In the case of product supply standards, the ultimate sanctions will be in the form of discontinuation of business-to-business transactions.

So, for ‘compliance control’ we look at the frequency of audits are performed, as well as the type of sanctions that are used. Without sanctions for non-compliance, it is not likely that the intended behavioural changes and intended environmental and socio-

economic impacts will occur. So, here we need to know if labelling will be terminated after established non-compliance.

Some SSCG-systems are more stringent than others, some allowing firm to adjust within a limited time frame if non-compliance occurs. Such an adjustment condition is already addressed under the performance component 'precision', so it will be ignored here.

In assessing compliance control, the more *frequent control* takes place, the higher its score (maximum of 1). The score for the sanctions is determined by looking at the *implementation time* for a adjustment preventing a sanction. A short timeframe to implement the requirements (score 1) will be getting a higher score than a longer timeframe (0.66 and 0.33 points). The two scores for *control frequency* and *implementation time* are multiplied.

3.5 Method outline

The method assesses and calculates scores for the different links in the supply chain and the different performance components. The outcome for coverage and precision is measured by using a scoring matrix. In order to calculate compliance control a different method is used. The outline of the scoring matrix is provided in Table 2. It shows the links of the coffee supply chain in the columns (chain coverage) and the sustainability aspects as explained in paragraph 3.2 in the rows (subject coverage). As most product standards do not include the supply chain links after transport to the consumer country, this is not included in this study³.

Box 1 Model summary

Coverage

C_{total} = Coverage total score

$$C_{total} = ((C_{soc} + C_{env} + C_{econ}) / 3) \times 100$$

Precision

P_{total} = Precision total score

$$P_{total} = ((P_{soc} + P_{env} + P_{econ}) / 3) \times 100$$

Compliance Control

CC_{total} = Compliance Control total score

$$CC_{total} = ((CC_{control} + CC_{sanctions}) / 2) \times 100$$

Governance System Performance

GSP = Governance System Performance

GSIP = Governance System Impact Potential

$$GSP = C_{total} \times P_{total} \times CC_{total}$$

$$GSIP = GSP \times \text{market share SSCG system}$$

The 'total coverage' (C_{total}) is found by looking at the presence of the sustainability aspects addressed at the different links of the supply chain. The presence of these sustainability aspects is found in the different requirements of the certification standard. A score of 1 is given in each box if the sustainability subject is addressed in the text of the standard. When a subject is not addressed, a score of 0 is given in the boxes. It is important to note that a score only can be given if something is *specifically addressed* in a written standard and can not be given based upon assumptions, contextual information, or vague descriptions. As there are a lot of requirements within a standard, it is very well possible that one cell of the matrix has multiple scores. When this is the case, the average of each cell needs to be calculated. The next step will be to measure the average for each sustainability sphere (environment, society, economy: C_{soc} ; C_{env} ; C_{econ}). Thus, for C_{soc} this implies that the score in the social sphere is divided by the maximum score of 25 ($25 = 5 \text{ aspects} \times 5 \text{ chain links}$) to find the average score. This needs to be done for each sustainability subject. Added together ($C_{total} = ((C_{soc} ; C_{env} ; C_{econ})/3) \times 100$), divided by 3 and multiplied by 100, it will give the final score in percentages. This score represents the percentage of the supply chain that is covered with sustainability aspects.

The 'total precision' is found by looking at the degree of precision in the same indicator groups. For each indicator group addressed under 'coverage', a score is given with 1 for a high requirement level, 0.66 for a low requirement level, 0.33 for a recommended requirement level and 0 for no precision indicated at all.

For each level (high, low and recommended), different keywords are identified to make sure that the different requirements in the certification are analysed on a proper basis.

For the 'high' requirement level, the key words are must, have (ought) to, ought to and obliged (to). Beside these keywords, there could also be a very precise time indication or specific numbers that are required.

For the 'low' requirement level, an important key word is should, but it is more important that the high requirement level keywords are not used in the requirement, that numbers are not given and a timeframe is not specifically mentioned.

For the 'recommended' level, the key words of the high and low requirement level are not present, words such as if, could and recommended are included in the requirements and the precision of this requirement is very vague. As it is possible that more than one indicator is present in one box of the matrix, the average score for each box needs to be calculated. With this calculation, a maximum score of 1 per box can be obtained.

So, for 'precision', a similar way of assessing the different requirements is done as for 'coverage'. For each requirement it is checked what the precision is. It must be noted that the precision can only be found for those requirements, which also have been assessed under 'coverage', as both are closely related. As with 'coverage', an average score for each sphere (P_{soc} , P_{env} , P_{econ}) is found by dividing the given score by the maximum score to be obtained in that sphere (which is 25 for P_{soc} , 30 for P_{env} and 20 for P_{econ}). The final score is determined by the following calculation: $P_{total} = ((P_{soc} + P_{env} + P_{econ}) / 3) \times 100$.

For 'compliance control' another method of calculation is used. As compliance control (CC_{total}) is about *control* and *sanctions*, both need to be measured. Here we used the established practice in auditing environmental management systems (ISO 14001, EMAS) as a reasonable reference. The control frequency can either be a half year, year, two year, or two year-plus basis, where half year basis is more positive than the two year-plus basis: more control will imply a better performance of the supply chain. The

‘*sanction*’ concept refers to whether there are sanctions for non-compliance and to the time given by the certification body to a producer/grower to fix the non-compliances with the certification scheme.

The *control* score is based on the interval between two control moments. The more frequent production sites and factories are being controlled, the higher the score, with a maximum score of 1 when control takes place on a regular and frequent basis. When the control takes place every half a year up to a year the score is 1, when the control takes place on a two year basis, the score is 0.66 and when the control takes place on a 2+ year basis, the score will be 0.33. This will be called CC_{control} .

Second, for the *sanctions* score, a maximum score of 1 can be obtained when there are sanctions on non-compliance. No sanctions will be getting a score of 0 (zero). When there are minor sanctions, but they have a time span for repair of over two years, a score of 0.33 will be given. A score of 0.66 will be given for somewhat stricter sanction where non-compliance needs to be solved within one year. A score of 1 will be given when non-compliance needs to be solved within half a year. This outcome will be called CC_{sanction} . The two outcomes of the control and the sanctions will be added together and divided by two another to get the CC_{total} score; $CC_{\text{total}} = ((CC_{\text{control}} + CC_{\text{sanctions}}) / 2) \times 100$. When the score is multiplied by 100, the percentage of the compliance is found.

The final step is to determine the GSIP. The GSIP is the product of the *GSP* (coverage, precision and compliance control) and the *market share* of the selected SSCG system:

$$GSIP = \text{market share} \times GSP.$$

The maximum score to be found with this multiplication is 1. The actual outcome of this calculation can well be expected to be low, as both the GSP and GSIP depend on the product of different components. This implies a curve linear relation with the product rising more strongly, when values of the individual components get closer to the maximum value of 1. And, the other way around, if one of the components has a very low score, the product score will also remain low. For instance, there may be a high score for ‘*coverage*’, but if the ‘*precision*’ is poorly addressed in the standard, the value of ‘*coverage*’ will not be effective. The same goes for ‘*compliance control*’. If all the other components are valued high, but there is no control or no sanction, it is not sure whether there actually is a performance to talk about, since it cannot be checked. By using this way of calculating, we assume to indeed represent a true impact potential, expressing that each component is a necessary, but in itself insufficient condition to create impact.

3.6 How to use the methodology

In the previous paragraphs the method was explained. This section explains how the method should be applied to the sustainability (certification) standards created in supply chain governance systems. The values for *coverage* and *precision* are to be determined for each indicator ‘cell’ in Table 2 (which represents an aspect at a specific link in the chain), based on the specific requirements formulated in the written standard (formulated as ‘articles’ in the list of requirements). For each cell, it is possible that there is only one relevant requirement, but also multiple requirements can be relevant. A score of 1 is given if there is any relevant requirement available for this cell. When all requirements in the standard relevant for a specific cell are scored, the average for each cell of ‘*coverage*’

needs to be calculated with a maximum score of 1. The value for '*precision*' is found by determining the precision of formulation of the requirements found in each cell where the '*coverage*' was assessed. As with '*coverage*', this implies that there can be multiple requirements, which are relevant for the indicator cell. When all requirements are assessed, the average '*precision*' score can be calculated for each cell with also a '*coverage*' score.

It is important that with the assessment of '*coverage*' and '*precision*' the calculated score is assigned to the right link of the supply chain, as not all requirements apply to every link of the supply chain. If all requirements are assessed with the associated indicators for '*coverage*' and '*precision*', and assigned to the right links in the supply chain, the average for each cell of the matrix can be determined.

Let us use as an example the 'poverty reduction' aspect in the year 2003 of the Fairtrade standards to illustrate this for '*coverage*'. For this aspect three indicator groups were identified (housing, fair income and human rights): in the entire text of the Fairtrade 2003 Standard, housing and human rights both were counted zero times, implying a score of zero. Fair income was counted in three specific articles, thereby scoring one (1), as it is just about the general presence of an indicator and not the precise number it appears in the text. The average is subsequently $(1 + 0 + 0) / 3 = 0.33$. This is done for each sustainability aspect. The overall scores for the sustainability aspects are subsequently summed to calculate the sustainability sphere average scores C_{soc} , C_{env} and C_{econ} . The same way of calculating the final score is used for '*precision*'.

For compliance control we do not use the scoring matrix, because it is not based on the standard document. To find the required results for compliance control, it is necessary to have interviews with the audit organisations and to find information in available documents about the compliance control practice. We do not need to specify this for all cells in the matrix, but can work with one value per standard for this performance component.

4 Case studies

In order to verify the method, the case studies chosen are two coffee SSCG systems dominant on the Dutch market. As discussed in Section 2 the emergence of sustainable coffee has a history of two decades, with more recently various competing certification systems and acknowledged sustainability initiatives in the market. Despite the international developments made in the international coffee industry, the cases that were chosen are both Dutch. The Dutch coffee market has one of the biggest market shares in the world regarding sustainable development with roasters such as Douwe Egberts/Sara Lee or De Drie Mollen Holding (Giovanucci and Koekoek, 2003). Furthermore, several of the existing SSCG systems originated in the Netherlands. Two of these systems have been chosen to study: Fairtrade Coffee ('Max Havelaar') and Utz Certified, as they represent the largest market shares in the Netherlands.

4.1 Fairtrade

Fairtrade is a strategy for poverty alleviation and sustainable development (FLO, 2009) with the aim to provide an 'alternative approach to conventional trade and improving livelihoods and well-being of small producers by improving their market access,

strengthening their organisations and paying a fair price and a continuous relationship' [Giovanucci and Koekoek, (2003), p.38]. It originated as a partnership between the Alternative Trading Organisations (ATOs) in the 1950s and 1960s. The labelling of the products did not start until 1988, with the Max Havelaar initiative for coffee in the Netherlands. This standard setting process evolved over the years and now includes an extensive process including stakeholder consultation, verification and compliance control. This Dutch initiative as well as other international Fairtrade Initiatives joined their forces and formed the Fairtrade Labelling Organisation (FLO): an umbrella organisation mandated to apply and monitor Fairtrade standards worldwide⁴. The role of the national branches of FLO, such as Max Havelaar is limited to licensing importers and distributors in their own country in accordance with the guidelines of the FLO (Raynolds, 2009).

Fairtrade labelled coffee is purchased directly from small farmer cooperatives which are located in approximately 30 different countries. This coffee is guaranteed to be produced under contractual social and environmental standards and is purchased at a minimum contract price which includes a social premium, which the producers are obliged to invest in the community, education, healthcare or infrastructure (Daviron and Ponte, 2005; Giovanucci and Koekoek, 2003). While this price previously could vary per region, the current (which dates from 2008) price on coffee is set at a worldwide level.

As soon as they have become a member, the farmers do not have to pay any membership fees, but need to comply to Fairtrade standards. While auditing and compliance control previously was carried out by FLO itself, it is now carried out by FLO-Certified (FLO-Cert): an autonomous certification body of FLO. The trader or the company responsible for applying the Fairtrade label are also audited by FLO-Cert to make sure that they comply to the standards as well. They on the other hand do need to pay a US\$ 0.10/lb fee for licensing (Potts et al., 2007).

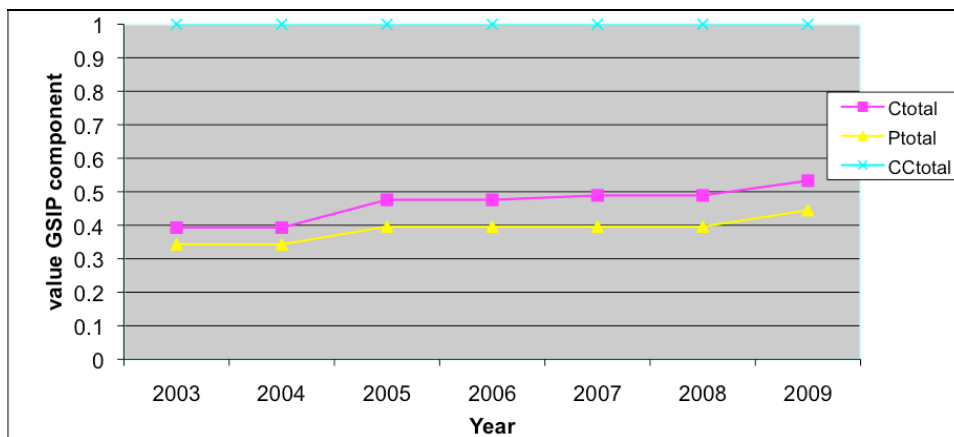
The Fairtrade standards have been revised over the years with minor and major changes to improve the standard and its output. The General Trade Standard is the standard that, as the name implies, is generally applicable to different kinds of commodities Fairtrade certifies. In addition to the General Trade Standards, Fairtrade has created specific trade standards for different commodities such as cotton, cocoa and coffee, with the latter coming into practice in 2003. The full name of the specific trade standard for coffee is 'Fairtrade Standard for coffee' and consist out of five parts:

- 1 an introduction to Fairtrade
- 2 serving as a guidance for the rest of the standards
- 3 part A: generic Fairtrade standards for small farmers' organisations (including categories as social development, economic development, environmental development and standards on labour conditions)
- 4 part B: product specific standards for coffee which does not contain any extra requirements to part A
- 5 part C: trade standards for coffee.

Over the years, the Generic Trade Standards (part A) have been revised in the year 2003, 2005, 2007 and 2009. While some years the changes were minor, other years major revisions on e.g. environmental issues were made. Part B and has not changed over the years and part C has undergone some minor changes.

For each year in which a revision has been made, the GSP components were measured for Fairtrade (Table 3). The results over the years are plotted in a line diagram, shown in Figure 3.

Figure 3 GSP components Fairtrade coffee standard (see online version for colours)



It shows that both the coverage (C_{total}) and precision (P_{total}) improved over the years. The growth in 2005 is the result of the incorporation of new environmental requirements. Before, Fairtrade did not have many environmental requirements, while in 2005, this increased considerably, resulting in a better coverage and precision. Another revision was made in 2009, as a result of the renewed standard, which includes more guidelines on how the standards should be interpreted. This new standard covers more topics and also increased on precision. The relatively low score of both precision and coverage relates to the fact that Fairtrade only applies to farmers and does not include other links of the supply chain. This implies that only the first three rows of the matrix are included in the analysis instead of a full coverage of the entire supply chain. As for the compliance control, the value remains the same (1), as out of interviews it appeared to be as such that FLO-Cert did regular audits and sanctions on the producers. One should note that this was stated as general policy, while no specific data on inspection frequencies were made available for this study.

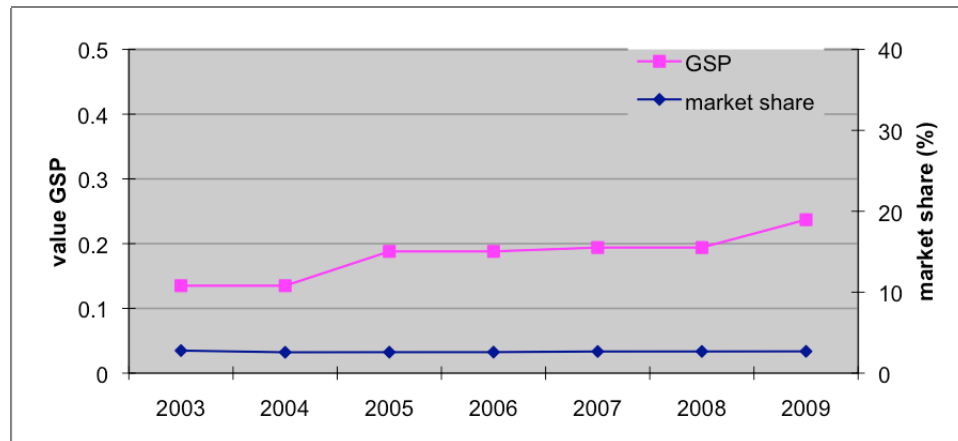
Table 3 GSP and market share Fairtrade coffee standard

Component\year	2003	2004	2005	2006	2007	2008	2009
C _{total}	0.393	0.393	0.476	0.476	0.489	0.489	0.533
P _{total}	0.343	0.343	0.396	0.396	0.396	0.396	0.445
CC _{total}	1	1	1	1	1	1	1
GSP	0.135	0.135	0.188	0.188	0.194	0.194	0.237
Market share in NL	2.8	2.6	2.6	2.6	2.7	2.7	2.7

The next is to calculate the GSP and then the GSIP, where we assume the impact to correlate with the market share (sales of certified divided by total sales in a national market). Finding the latter however is difficult. Even though we were able to find sales data of Fairtrade coffee in the Netherlands, we had to derive the total volume of the

market from indirect data. In this way we can show the GSP and the market share in Figure 4. As we can expect, the GSP line shows a picture comparable to the GSP components, slowly increasing from 0.135 to 0.237 in the period 2003–2009. This GSP shows the improved performance of the system. We also see hardly any increase in the small market share of Fairtrade coffee being 2.6 to 2.8% the whole period.

Figure 4 GSP and market shares Fairtrade coffee (see online version for colours)



4.2 *Utz Certified*

Utz Kapeh (it changed its name to Utz Certified in 2007) originated in 1999 but the foundation itself was not established until 2002. The ambition of Utz Certified is to “enable coffee producers and brands to credibility and transparently demonstrate their commitment to sustainability in a market driven way” (IISD, 2004). While originated in Guatemala and the Netherlands, the scope of Utz Certified broadened and the coffee is currently produced by producers located in approximately 21 countries in 2010 (Utz Certified, 2010).

The standard was based on the criteria of the European Retailers Produce group of Good Agricultural Practices (EUREP-Gap) established in 1997 and was officially recognised by this standard in 2004. In addition to these criteria, Utz Certified added extra criteria based on the ILO convention and the International Declaration of Human Rights. The 2009 standard has been completely separated from the EUREP-Gap standard and it was decided to create an independent standard reflecting on Utz Certified ambitions and ideas. Besides the Code of Conduct which applies to (an association of) farmers and/or producers, Utz Certified also pursues Chain of Custody requirements which apply to the importers and roasters, ensuring the quality of the Utz Certified product throughout the entire chain. Both the Chain of Custody and the Code of Conduct are renewed every few years.

If producers want to be certified by Utz Certified, they need to comply to the Code of Conduct. When a producer complies with these criteria, the Utz Certified certificate is granted (Ponte, 2004). Once a member, the certificate is renewed every year after an independent audit. The Utz Certified membership for both producers and

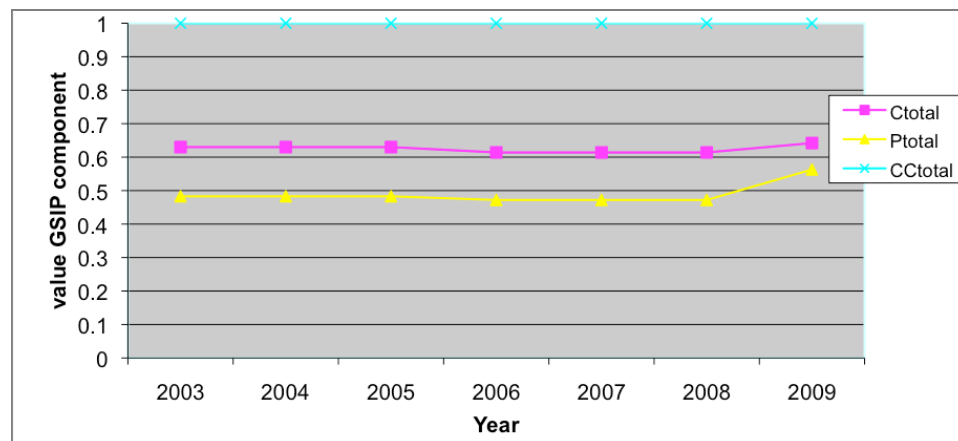
importers/roasters is not free. To stay a member, the farmer has to pay the organisation a fee of US\$ 0.02/lb of coffee.

Furthermore, Utz Certified does not offer a fixed premium on the regular coffee price. Instead, the roaster buys its coffee directly from the certified supplier and both supplier and roaster determine the right price for the coffee together. In practice this often results in producers getting a premium of between US\$ 0.07 (low) and US\$ 0.26 (high)/lb for Arabica coffee. Also the system works with a premium rewarding sustainability. The total price of Utz Certified coffee thereby consists of the market reference price (for Arabica or Robusta coffee) plus or minus a quality differential plus a variable sustainability premium (Daviron and Ponte, 2005). Despite of what the name 'premium system' might imply, Utz Certified does not interfere with these direct price negotiations between two market actors.

Both the Code of Conduct and the Chain of Custody have been revised three times in 2003, 2006 and 2009. The Code of Conduct focuses on three broad categories: social responsibility, environmental responsibly and food safety (Ponte, 2004). These three categories address issues on soil management, fertiliser use, integrated pest management, waste pollution management, health of the workers, safety and welfare and cultural health. While the current version of the Code of Conduct has only mandatory requirements which the farmer has to comply to within one to five years after being certified, older versions of the Utz Code of Conduct used a different system with different mandatory requirements, referring to a minor or low recommendation, implying that the producer did not have to comply to all criteria in order to become certified. The Utz Code of Conduct applies to all coffee producers producing and selling coffee as Utz Certified; this can either be individual farms or a group of organised producers (Utz Certified, 2009). The Chain of Custody is required for each party who has legal ownership of Utz Certified coffee (Utz Kapeh, 2005).

For each of the years when changes were made, the GSP components were measured for Utz Certified (Table 4). The corresponding values are plotted in a line diagram in Figure 5.

Figure 5 GSP components Utz Certified standard (see online version for colours)



Overall, Utz Certified has a higher value than Fairtrade for coverage and precision. This is a result of the broader scope of the standard with both the Chain of Custody and Code

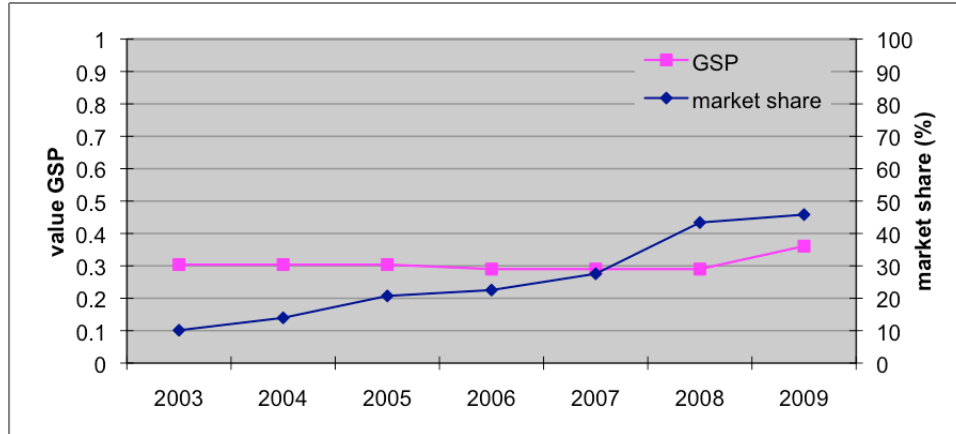
of Conduct covering each link in the supply chain. We also see a slight drop in coverage and precision in 2006, with an increase in 2009. The slight dip in 2006 is the result of integrating two requirements into one, thereby decreasing the number of requirements, which could be valued. In 2009 the coverage increased slightly, while precision increased more. This is a result of two developments. First of all, the score of the Utz Certified standard increased due to adding more requirements on social and environmental development. Secondly Utz Certified started to apply more stringent requirements in their standard, while previously they made a distinction with major, minor and recommended requirements. In the earlier standard, the certificate holder had to comply with 100% of the major control points, 95% of the minor control points and none of the recommendations in order to obtain a certificate. The 2009 standard made all requirements obligatory after either one or four years, implying that all requirements had to be taken into account. Previously this was not possible and farmers could choose to not (yet) comply with some requirements.

The practice of compliance control for Utz Certified has remained the same. The auditing for Utz Certified is arranged by an independent audit agency, which performs audits on a regular basis. Originally it was seen that there is very little non-compliance with Utz Certified, as a result of the non-compellingly and specific formulated requirements of Utz Certified. Depending on the type of non-compliance and the time it would rationally take to solve it, sanctions are provided. When problems are not solved after a short while, the certificate will be withdrawn.⁵ This results in a score of 1 from the years 2003 until 2009. Based on the non-compellingly and specific formulated requirements of Utz, it may be questioned why the precision score is not much higher. This is a result of the fact that before 2009, Utz offered the farmers the possibility of non-compliance to some of the standards. The requirements were therefore precise in what had to be done, but not when it had to be done. From the 2009 standard onwards, this became more precise, which is why the scoring for precision in 2009 is higher than in the previous years.

Table 4 GSP and market share Utz Certified standard

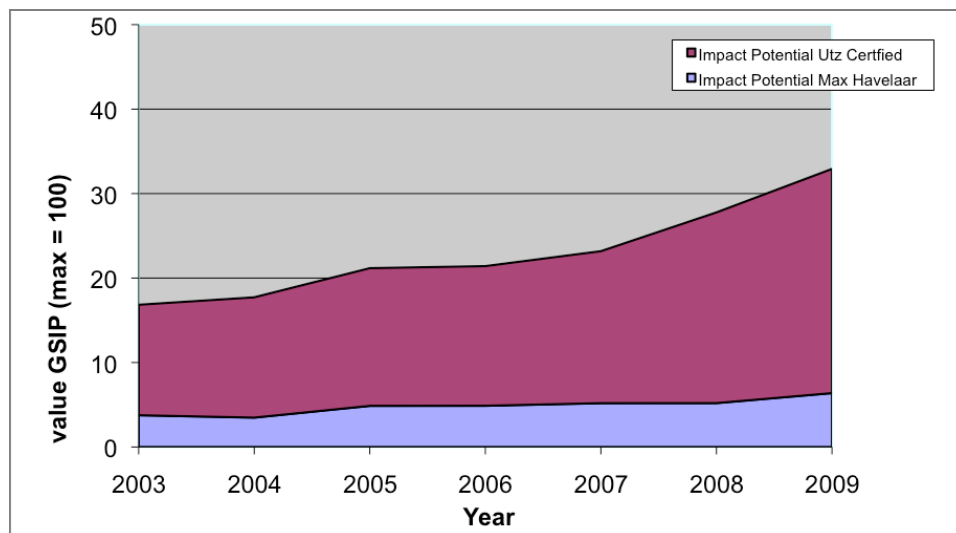
<i>Component\year</i>	<i>2003</i>	<i>2004</i>	<i>2005</i>	<i>2006</i>	<i>2007</i>	<i>2009</i>	<i>2009</i>
C _{total}	0.630	0.630	0.630	0.614	0.614	0.614	0.642
P _{total}	0.483	0.483	0.483	0.472	0.472	0.472	0.563
CC _{total}	1	1	1	1	1	1	1
GSP	0.304	0.304	0.304	0.290	0.290	0.290	0.361
Market share in NL	10.0	14.0	20.7	22.5	27.6	43.4	45.8

As with Fairtrade, the final step is to calculate the GSP and GSIP. Again, it was difficult to find sufficient data enabling the calculation of the GSIP, as it was even more difficult to obtain sales data for Utz Certified over the full period in the Netherlands. Utz Certified gave market share data for the Dutch market, starting from 2003. We can see the results in Figure 6. It shows a similar stable line as for different GSP components, with an improvement in 2009. The market shares for Utz Certified have been growing far faster than for Fairtrade, due to the recognition by a number of main coffee sellers and retailers on the Dutch market. The fastest growth was in the period before 2008, and in 2009 the market share grew to 47%.

Figure 6 GSP and market shares Utz Certified (see online version for colours)

4.3 Combined impact potential

With these data we can now show the accumulated impact potential of both systems in time. This is illustrated in Figure 7. We see that a stable fundament in the market was created by the first system, Fairtrade. Combining the GSP with the market share its value is 6.38 on a scale to 100. The GISP values for Utz Certified increase to the level of 35.32. The combined impact potentials of Fairtrade and Utz Certified have thus grown from 16.9 to 41.7, which is a 147% improvement in the period 2003–2009. Before 2003 Utz Certified was not yet operational, which implies an even larger improvement if we compare to the period before 2002 (up to 12 times larger).

Figure 7 GISP max Havelaar and Utz Certified (see online version for colours)

5 Discussion and conclusions

What appears from the pilot case studies is that the method is feasible and is capable of clearly illustrating the developments in the full market. We also see that the existing systems do not yet address the full potentials of sustainable production (coverage) or are not addressing all links at the supply of the product chain. It enables analysing in time and comparing between systems and between performance components.

One of the main weaknesses of the methodology is the reliance on the availability of market data and more accurate data on the practice of compliance control. Also the availability of previous versions of standards has been an issue, being difficult to get on the table during the research project. This in itself is a remarkable observation, where one would expect forms of self-governance in response to market demands to be more transparent

The case studies selected for testing the method were from the Netherlands for practical reasons. Most private standards for sustainable products still cover a very small portion of markets. The two Dutch cases had an advantage because of the recent market break through. For reasons of practical limitations no attempts were made to look at other countries. As the GSIP only can be determined with sufficient market data, it might very well be that the GSIP method can be easier applied in other countries if better market data is available.

The first part of the method assessing the GSP components and thus calculating the GSP in our opinion shows a realistic representation of the reality. We must consider that it will be hard for a certification system to obtain a perfect score of 100% on the GSP indicator. The outcomes of the case studies show that Fairtrade, with a limited scope in terms of covering all links, scores lower than Utz Certified, mainly because it covers more links of the supply chain. Furthermore, major and minor changes, which have been made in the standard over the years, can be traced back in the outcomes. In particular 'precision' seems to be more sensitive to these changes, as it takes a more precise look at the requirements. The best example for this is the Utz Certified 2006 standard, which integrated several requirements into one major requirement. This is well reflected in the final outcomes of the GSP method.

Measuring 'compliance control' calls for some discussion. While the audit organisation of Utz Certified was very open and honest about the outcomes for Utz Certified, it was difficult to find the right information also for Fairtrade. The information used in the method was in the end retrieved from an electronic interview. To find the exact information, it is needed that supply chain governance systems become more transparent about their performance. One has to keep in mind that the type of information required for this method is not just only relevant for academic purposes. In fact, it is information relevant for any of the stakeholders connected to supply chain governance: the involved supply chain partners, including the final consumer of the product, the sector organisations and non-governmental organisations and also governments that need to track the progress made in these self-regulating markets. However, first of all these newly formed certification organisations in the global competition for delivering sustainable products, themselves should be interested in these key performance indicators and willing to publish them on their websites.

Combining the four components equally in the calculation of the GSIP implicitly weighs the included elements equally. The same goes for the components in the coverage aspect. This may call for some debate, however for the purpose of testing the

methodology we did not develop any line of reasoning into specified weighing of the elements. One line of reasoning here might be to add a relative weight to either the various aspect coverage components or the chain coverage components in Table 2. This would only make sense if it could be based on detailed life cycle analysis, which would open the doors to a wide variety of place and time diversity in LCA outcomes. This would probably not result in any consensus and thus not yield a better result in assessing the 'potential' impact of a certification system.

Yet, adding weighing factors to the formulas, can easily be done, and could also serve other goals. That is if users of the system would want to adopt the outcomes to their preferences, for example if a user (a firm, a consumer, or a consumer organisation) prefers to focus on just some of the sustainability aspects or focus only on the first links in the supply chain.

With the data applied here for illustrating the methodology we have sufficiently shown the reliability and applicability of the method. With further detailing and more accurate compliance and market share data the outcomes might only become more accurate and represent the actual situation even better.

We expect that further elaboration of the methodology can serve various needs. Managers of certification systems can use it to benchmark and improve their own system. Users in the supply chain can benefit from it by identifying the most suitable certification system to work with and identify opportunities for improvement in their own and previous link of their supply chain. Consumers and their organisations can apply it as a benchmark tools and show results in the information to consumers about the quality of the overwhelming supply of logo's and labels (Vermeulen and Kok, 2012). Finally, governments can use it for monitoring the quality of private certification systems and use a certain minimum threshold as requirement for eligibility for various kinds of government incentives (such as public procurement policies).

References

- Albersmeier, F., Schulze, H. and Spiller, A. (2009) 'Evaluation and reliability of the organic certification system: perceptions by farmers in Latin America', *Sustainable Development*, Vol. 17, No. 5, pp.311–324.
- Alvarez, G., Pilbeam, C. and Wilding, R. (2010) 'Nestlé Nespresso AAA sustainable quality program: an investigation into the governance dynamics in a multi-stakeholder supply chain network', *Supply Chain Management*, Vol. 15, No. 2, pp.165–182.
- Chan, F.T.S. and Qi, H.J. (2003) 'An innovative performance measurement method for supply chain management', *Supply Chain Management*, Vol. 8, No. 3, pp.209–223.
- Ciliberti, F., Pontrandolfo, P. and Scozzi, B. (2008) 'Investigating corporate social responsibility in supply chains: a SME perspective', *Journal of Cleaner Production*, Vol. 16, No. 15, pp.1579–1588.
- Coffee Coalition (2007) *Coffee Barometer 2006: Certified Coffee in the Netherlands*, The Coffee Coalition, The Hague.
- Consumers International and IIED (2005) *From Bean to Cup: How Consumer Choice Impacts on Coffee Producers and the Environment*, Consumer International, London.
- Costanza, R. and Daly, H.E. (1992) 'Natural capital and sustainable development', *Conservation Biology*, Vol. 6, No. 1, pp.37–46.
- Daly, H.E. (1990) 'Towards some operational principles of sustainable development', *Ecological Economics*, Vol. 2, No. 1, pp.1–5.

- Daviron, B. and Ponte, S. (2005) *The Coffee Paradox: Global Markets, Commodity Trade and The Elusive Promise of Development*, Zed Books, London and New York.
- FLO – Fairtrade Labelling Organisation (2009) *What is Fairtrade* [online] http://www.fairtrade.net/what_is_fairtrade.html (accessed 18 December 2009).
- Gillison, A.N., Liswanti, N., Budidarsono, S., Van Noordwijk, M. and Tomich, T.P. (2004) 'Report: impact of cropping methods on biodiversity in coffee agro-ecosystems in Sumatra, Indonesia', *Ecology and Society*, Vol. 9, No. 2, pp.7–23.
- Giovannucci, D. and Purcell, T. (2008) *Standards and Agricultural Trade in Asia*, ADBI Discussion Paper 107, Asian Development Bank Institute, Tokyo [online] <http://www.adbi.org/discussion-paper/2008/05/22/2542.standards.agricultural.trade.asia/AsianDevelopment> (accessed 17 September 2014).
- Giovanucci, D. and Koekoek, F.J. (2003) *The State of Sustainable Coffee: A Study of 12 Major Markets*, ICO and IISD, London and Winnipeg.
- Gopal, G. and Cline, S. (2007) 'Driving towards sustainable profitability: transportation service providers and customer relationship management', *Supply Chain Management*, Vol. 12, No. 2, pp.85–87.
- Green, K., Morton, B. and New, S. (1998) 'Green purchasing and supply policies: do they improve companies' environmental performance?', *Supply Chain Management*, Vol. 3, No. 2, pp.89–95.
- IISD – International Institute for Sustainable Development (2004) *Environmental and Other Labelling of Coffee. The Role of Mutual Recognition, Supporting Cooperative Action*, IISD, Canada.
- Iles, A. (2007) 'Making the seafood industry more sustainable: creating production chain transparency and accountability', *Journal of Cleaner Production*, Vol. 15, No. 6, pp.577–589.
- Keating, B., Quazi, A., Kriz, A. and Colman, T. (2008) 'In pursuit of a sustainable supply chain: Insights from Westpac Banking Corporation', *Supply Chain Management*, Vol. 13, No. 3, pp.175–179.
- Kolk, A. (2005) 'Corporate social responsibility in the coffee sector: the dynamics of MNC responses and code development', *European Management Journal*, Vol. 23, No. 2, pp.228–236.
- Kovács, G. (2008) 'Corporate environmental responsibility in the supply chain', *Journal of Cleaner Production*, Vol. 16, No. 15, pp.1571–1578.
- Lozano, R. and Huisingh, D. (2011) 'Inter-linking issues and dimensions in sustainability reporting', *Journal of Cleaner Production*, Vol. 19, Nos. 2–3, pp.99–107.
- McIntyre, K., Smith, H., Henham, A. and Pretlove, J. (1998) 'Environmental performance indicators for integrated supply chains: the case of Xerox Ltd', *Supply Chain Management*, Vol. 3, No. 3, pp.149–156.
- Metselaar, J.A. (2010) *The Sustainable Supply Chain Governance (SSCG) System Performance. A Longitudinal Analysis*, Master thesis. Utrecht University, Utrecht.
- Ponte, S. (2004) *Standards and Sustainability in the Coffee Sector – A Global Value Chain Approach*, International Institute for Sustainable Development, Winnipeg.
- Potts, J., Fernandez, G. and Wunderlich, C. (2007) *Trading Practices for a Sustainable Coffee Sector. Context, Strategies and Recommendations for Action*, IISD – International Institute for Sustainable Development, Canada.
- Ras, P.J. and Vermeulen, W.J.V. et al. (2007) 'Greening global product chains: bridging barriers in the north-south cooperation. An exploratory study of possibilities for improvement in the product chains of table grape and wine connecting South Africa and the Netherlands', *Progress in Industrial Ecology – An International Journal*, Vol. 4, No. 6, pp.401–417.
- Raynolds, L.T. (2009) 'Mainstreaming fair trade coffee: from partnership to traceability', *World Development*, Vol. 37, No. 6, pp.1083–1093.

- Schaltegger, S. (2002) 'A framework for ecopreneurship: leading bioneers and environmental managers to ecopreneurship', *Greener Management International*, Summer, No. 38, pp.45–58.
- Schliephake, K., Stevens, G. and Clay, S. (2009) 'Making resources work more efficiently – the importance of supply chain partnerships', *Journal of Cleaner Production*, Vol. 17, No. 14, pp.1257–1263.
- Seuring, S. and Müller, M. (2008) 'From a literature review to a conceptual framework for sustainable supply chain management', *Journal of Cleaner Production*, Vol. 16, No. 15, pp.1699–1710.
- Seuring, S., Sarkis, J., Müller, M. and Rao, P. (2008) 'Sustainability and supply chain management – an introduction to the special issue', *Journal of Cleaner Production*, Vol. 16, No. 15, pp.1545–1551.
- Svensson, G. (2007) 'Aspects of sustainable supply chain management (SSCM): conceptual framework and empirical example', *Supply Chain Management*, Vol. 12, No. 4, pp.262–266.
- TCC (2009) *Coffee Barometer*, S. Panhuysen and B. Weiligmann, Tropical Commodity Coalition, The Hague.
- Testa, F. and Iraldo, F. (2010) 'Shadows and lights of GSCM (green supply chain management): determinants and effects of these practices based on a multi-national study', *Journal of Cleaner Production*, Vol. 18, Nos. 10–11, pp.953–962.
- Tsoufias, G.T. and Pappis, C.P. (2008) 'A model for supply chains environmental performance analysis and decision making', *Journal of Cleaner Production*, Vol. 16, No. 15, pp.1647–1657.
- United Nations Department for Sustainable Development (UNSD) (2007) *Indicators for Sustainable Development. Guidelines and Methodologies*, 3rd ed., October, United Nations, New York.
- United Nations Statistics Division (1997) *Glossary of Environment Statistics, Studies in Methods*, Series F, No. 67, United Nations, New York.
- Utz Certified (2009) *Utz Certified Good Inside Code of Conduct*, January, Utz Certified Good Inside, Netherlands.
- Utz Certified (2010) *Certified Producers* [online] <http://www.Utzcertified.org/> (accessed 9 March 2010).
- Utz Kapeh (2005) *Utz Kapeh Chain of Custody Requirements for Destination Countries*, Utz Kaph Foundation, Netherlands.
- Vermeulen, W.J. (2013) 'Self-governance for sustainable global supply chains: can it deliver the impacts needed?', *Business Strategy and the Environment*, Article online, DOI: 10.1002/bse.1804.
- Vermeulen, W.J.V. (2010) 'Sustainable supply chain governance systems: conditions for effective market based governance in global trade', *Progress in Industrial Ecology – An International Journal*, Vol. 7, No. 2, pp.138–162.
- Vermeulen, W.J.V. and Kok, M.T.J. (2012) 'Government interventions in sustainable supply chain governance: experience in Dutch front-running cases', *Ecological Economics*, November, Vol. 83, pp.183–196.
- Vermeulen, W.J.V. and Ras, P. (2006) 'The challenge of greening global product chains: meeting both ends', *Sustainable Development*, Vol. 14, No. 4, pp.245–256.
- Vermeulen, W.J.V. and Seuring, S. (2009) 'Introduction: sustainability through the market: the impacts of sustainable supply chain management', *Sustainable Development*, Vol. 17, No. 5, pp.245–256.
- Vermeulen, W.J.V. and Uitenboogaart, Y. et al. (2010) *Roles of Governments in Multi-Actor Sustainable Supply Chain Governance Systems and Effectiveness of their Interventions: An Exploratory Study*, Netherlands Environmental Assessment Agency (PBL) and Utrecht University, Bilthoven.
- Williamson, O.E. (2008) 'Outsourcing: transaction cost economics and supply chain management', *Journal of Supply Chain Management*, Vol. 44, No. 2, pp.5–16.

Youn, S., Yang, M.G., Hong, P. and Park, K. (2013) 'Strategic supply chain partnership, environmental supply chain management practices, and performance outcomes: an empirical study of Korean firms', *Journal of Cleaner Production*, 1 October, Vol. 56, pp.121–130.

Notes

- 1 This article is based on a study for the Dutch Environmental Assessment Agency PBL, called "The sustainable supply chain governance (SSCG) system performance: a longitudinal analysis" by Metselaar (2010).
- 2 See also Metselaar (2010) for a more elaborated explanation on how sustainable development is defined.
- 3 This pragmatic choice can actually be debated. In practice these private standards are mostly developed in developed countries and one can argue that most of these issues are well addressed in national policies in these countries and do not need to be included in standards. Working at the supply side as we often experience strong opposition from market actors to this, arguing that they also want to be assured of fair practices at European importers, producers and retailers (e.g., Ras et al., 2007). The methodology would actually easily be adapted to accommodate such demands. For the purpose of illustration here we focus in our methodology on the measuring only at the supply side.
- 4 Interview with N. Roozen, October 2009, Solodaridad, Utrecht, the Netherlands.
- 5 Interview with Control Union, October 2009; Interview with Utz Certified, November 2009.