Do Codes of Conduct Improve Worker Rights in Supply Chains?

A Study of Fair Wear Foundation

Niklas Egels-Zandén, PhD
Associate Professor
Department of Business Administration
School of Business, Economics and Law
University of Gothenburg
Box 600
SE – 405 30 Gothenburg, Sweden
niklas.egels-zanden@handels.gu.se
+46-31-7862729 (telephone)

Henrik Lindholm
Department of Philosophy and the History of Technology
Royal Institute of Technology
Teknikringen 78 B
SE – 100 44 Stockholm, Sweden.
helindho@kth.se
+46-706-473674 (telephone)

Forthcoming: *Journal of Cleaner Production*

Do Codes of Conduct Improve Worker Rights in Supply Chains?
A Study of Fair Wear Foundation

Abstract

The rise of private regulation of sustainability in global production networks has led to intensive debates about the impact of this regulation at the point of production. Yet, few empirical studies have systematically examined this impact in practice. Based on multiple factory audits of 43 garment factories conducted by the multi-stakeholder initiative Fair Wear Foundation, we show that codes of conduct improve (although marginally) worker rights on an overall level but that few significant results are found for specific worker rights. Our findings also lend support to the widespread argument that codes have uneven impact. Furthermore, we show that even rigorous multi-stakeholder factory audits are unable to identify process rights violations (such as those affecting freedom of association and discrimination), and that auditing is thus is more fundamentally flawed than assumed in previous research. Given companies' extensive investments in private regulation of worker rights, the findings have important implications for both scholars and managers.

Key words: code of conduct, garment, private regulation, supply chain, value chain, worker rights

Word count: 9538

1. Introduction

Trade is increasingly conducted via global production networks involving interlinked suppliers coordinated by lead companies that exert control without formal ownership. The rise of global production networks facilitates the global integration of activities from initial design to inputs, manufacturing, and distribution, through to the final retailing of goods and services (Barrientos et al., 2011b). It has also led to a fragmentation and geographical dispersion of production, typically in low-wage developing countries (Locke et al., 2013; Seuring and Müller, 2008).

This transformation of global production has created both opportunities and challenges for workers. On the one hand, the expansion of global production in labor-intensive industries has been an important source of employment generation, especially for women and migrant workers who previously had difficulty accessing wage employment (Barrientos et al., 2011a). On the other hand, stiff competition among export manufacturers in developing countries have led to poor working conditions and lax environmental standards in the factories producing for global brands (Locke et al., 2013). National governments have been unable to address these sustainability challenges, leading to multinational corporations being required, via voluntary private regulatory systems, to enforce sustainability principles at their legally independent and geographically dispersed suppliers (Barrientos et al., 2011b; Seuring and Müller, 2008). This rise of "private regulation" of worker rights followed from high-profile, activist-driven "name-and-shame" campaigns and has mainly taken the form of companies adopting codes of conduct and auditing schemes across their international network of suppliers (Bartley, 2007; Locke et al., 2007a).

With companies, labor activists and scholars having, since the early 1990s, invested staff, time, and resources into codes of conduct and auditing, the central question now becomes whether working conditions at the factories have improved, thanks to the codes. This question is intensely debated, with some scholars claiming that codes hold great promise for improving worker rights (e.g., Pearson and Seyfang, 2001; Ruggie, 2004; Zadek, 2004), while others claim that codes are too weak for the job (e.g., Blowfield and Dolan, 2008; Frundt, 2004).

The scholarly conversation about the effectiveness of codes of conduct is limited in three important ways. First, studies of the impact of codes have almost exclusively been limited to

qualitative studies of a small number of firms (e.g., Chan and Siu, 2010; Locke et al., 2013) and these should be complemented with more quantitative research. Second, the few quantitative studies have relied on corporate auditing methods (Locke et al., 2007a, 2007b) that previous research has identified as suffering from quality problems (Egels-Zandén, 2007; O'Rourke, 1997). Third, despite this lack of studies involving larger samples of suppliers, the view is widespread that codes have uneven impact, i.e., codes are presented as improving *outcome* standards but not *process* rights (Barrientos and Smith, 2007).

We seek to close these gaps by asking: *Do codes of conduct improve working conditions (in the form of outcome standards and/or process rights) at the point of production?* We draw on multiple factory audits of 43 garment factories that were conducted by Fair Wear Foundation (FWF). FWF is an independent non-profit organization that is recognized as conducting some of the most rigorous and trustworthy audits worldwide. In this way, we provide one of the first large-scale studies of code effectiveness based on credible auditing data. Through detailed coding of the factory audits, we also provide the first large-scale analysis of the claim that codes improve outcome standards but not process rights. Our results contribute to the scholarly conversation by showing that even rigorous factory audits are unable to identify violations of process rights; that codes do improve suppliers' overall performance, although marginally; that few significant results are found in relation to specific worker rights; and that one of the main problems with codes is that they are unable to ensure that compliant factories remain compliant over time.

2. Impact of codes of conduct on worker rights

The emergence of voluntary private regulation of worker rights has spurred an extensive literature on codes of conduct. In similarity to the literature into sustainable supply chains more generally (Seuring et al., 2008; Seuring and Gold, 2013), however, relatively few studies have empirically studied the impact of codes of conduct at the point of production in developing countries. As Seuring and Gold (2013: 1) put it, "among future research needs, supply chains in low-income countries stand out, which are still hardly addressed." Although incomplete, the existing body of research suggests two important insights about the impact of codes at the point of production. First, it highlights the challenges of ensuring that codes improve worker rights at the point of production where they are supposed to matter. For example, Chan and Siu (2010: 185), on the basis of a study of nine Chinese Wal-Mart

suppliers, show that the "general failure of auditing to detect violations of vital labor standards means that the CSR program of which Wal-Mart boasts has had little impact on workers at the company's supplier factories." Similarly, Egels-Zandén's (2007) study of Chinese suppliers to proactive Swedish toy firms found no factories (out of nine studied) in compliance with the Swedish firms' codes of conduct, with two-thirds of the suppliers violating all but one of the studied criteria.

While several other studies provide similar qualitative case evidence that "there has been little progress in improving labour standards through such [private] regulation" (Wells, 2007: 53), few studies move beyond individual cases. In a rare exception, Locke et al. (2007a, 2007b) used Nike's internal rating of 800 factories in 51 countries (audited between 1998 and 2005) and showed that, over time, almost half of the factories did not improve their compliance; 36% actually experienced a decline in compliance, and only approximately 20% improved. This was so even though Nike is recognized as investing extensively in code of conduct audits and is described as a code of conduct front-runner. Scholars in both qualitative and quantitative studies have thus shown that even when companies invest staff, time, and resources in codes and auditing, it is far from certain that the investment translates into improvements at the point of production.

At the same time, scholars comparing worker rights compliance by suppliers that are exposed or not exposed to codes have shown that those exposed to codes are generally more compliant. For example, Oka (2010a, 2010b) showed, on the basis of ILO-conducted audits through the Better Factory Cambodia project, that suppliers to reputation-conscious buyers that adopted strict codes and auditing were more compliant than suppliers to less reputation-conscious buyers. Similarly, Chakrabarty and Grote (2009) found from a survey in 2005 in India and Pakistan that child labor was less likely to be used in producing socially labeled than unlabeled carpets. It seems, then, as if suppliers exposed to codes and auditing are more compliant, but that it is questionable whether they improve over time as a consequence of codes and auditing.

Second, the research indicates that codes have uneven impact improving some, but not other, worker rights. Barrientos and Smith's (2007) argument has been particularly influential, claiming that codes improve *outcome* standards (such as health and safety, working times, and wages), while largely failing to improve *process* rights (such as trade union rights and

discrimination). This argument has been supported by several case studies of a small number of suppliers (Egels-Zandén, 2013; Frenkel, 2001; Mamic, 2004). The argument is *not* that codes are unable to improve process rights in specific cases (Rodríguez-Garavito, 2005; Ross, 2006), but simply that such improvements are unlikely. While the assumption of codes' uneven impact is widely disseminated, it has never been systematically tested.

To project these insights into the impact of codes further forward requires us to address five methodological shortcomings in previous research. First, it requires a move beyond the dominance of intensive case studies of small numbers of suppliers. There are signs of this development in recently published studies: for example, Locke et al. (2013) studied 276 HP suppliers, Anner (2012) studied 805 factory audits conducted by the Fair Labor Association (FLA), and Toffel et al. (2012) studied 31,915 audits conducted by a private auditing company. However, all these studies only compared *adherence* between different group of suppliers (i.e., the degree of supplier compliance with codes) and not *improvements* (i.e., comparing compliance over time). Similarly, Oka (2010a, 2010b) and Chakrabarty and Grote (2009) compared adherence (rather than improvements) between suppliers exposed and not exposed to codes. Locke et al.'s (2007a, 2007b) study of Nike currently remains the only quantitative study focusing on improvements over time.

Second, taking existing insights further requires abandoning the reliance on data collected at single points in time. Most qualitative studies of code impact rely on retrospective interview accounts to create reference points for which current working conditions are compared (e.g., Barrientos and Smith, 2007; Chan and Siu, 2010). This is problematic because of both high turnover in the studied industries (mainly garment) and the established research design weaknesses of reliance on retrospective accounts (e.g., Boring, 1954; Stouffer, 1949). In practice, this means that quantitative (and qualitative) studies of codes' impact must be based on, at least, two systematic audits of worker rights at a specific factory.

Third, the necessary research requires access to reliable factory audits since numerous studies have shown the weaknesses of factory audits of suppliers (Egels-Zandén, 2007; O'Rourke 1997). These weaknesses are particularly prevalent when audits are conducted by "commercial actors" (Barrientos and Smith, 2007) such as corporate internal auditors, specialized auditing firms (e.g., Intertek, Societé Générale de Surveillance, and Bureau Veritas) or service-driven, semi-commercial NGOs (Armbruster-Sandoval, 2005; Brown,

2013). Many of the audits conducted by commercial actors, for example, do not even include the central aspects of off-site interviews with workers and labor union representatives (Anner, 2012; Auret and Barrientos, 2004), leading auditors to rely on "data provided by management with little or no 'triangulation' or cross-checking" (Auret and Barrientos, 2004: 5). Given factory managements' tendencies to falsify records and instruct workers in how to respond when being audited, this is highly problematic (Egels-Zandén, 2007; Jiang, 2009; Taylor, 2011). These auditing problems raises doubts about the reliability of Locke et al.'s (2007a, 2007b) Nike study, since they relied on Nike's internally conducted audits, which, as the authors themselves recognized, could be biased.

Fourth, to assess the potentially uneven impact of codes requires access to detailed factory audits. This is often the case in qualitative studies of small numbers of suppliers (Barrientos and Smith, 2007; Chan and Siu, 2010; Egels-Zandén, 2007), but is lacking in most quantitative studies (see Brown et al., 2013a, 2013b, for an exception). For example, Locke et al. (2007a, 2007b) solely analyze a Nike-compiled aggregate percentage score for each supplier. This focus on aggregate measures resembles the focus in earlier quantitative studies on corporate social performance (CSP) and financial performance (FP), which often aggregated multiple CSP dimensions into an aggregate (e.g., Rowley and Berman, 2000). As Rowley and Berman (2000) argue, this leads to loss of richness and meaning in the data, questions of weighting of different dimensions, difficulties in comparison across studies, and results that are hard to interpret. Research into the impact of codes of conduct, thus, need to be based on empirical data that allow for analysis of both impact on a specific issue (such as child labor, working time, and trade union rights) and as an aggregate impact.

Finally, further advance in this research requires the ability to trace the development of a specific factory over time. While Locke et al. (2007a, 2007b) were able to analyze the aggregate performance of suppliers at a first and second audit in general, they were unable to trace whether suppliers' failure to improve was due to the codes being unable to turn non-compliance into compliance or due to new violations in previously compliant areas nullifying such improvements. In other words, is the main problem with codes that they are unable to achieve improvements, or that they are unable to sustain compliance over time?

In this paper, we start to address these five methodological shortcomings in previous research by undertaking a study of factory audits conduct by FWF. Table 1 summarizes the main differences between previous studies and the here presented FWF study.

Insert Table 1 about here

In other words, our study is one of the first quantitative studies based on reliable, detailed, and traceable empirical data on the impact of codes of conduct, and thanks to this it can provide the first quantitative analysis of the claim that codes have uneven impact.

3. Fair Wear Foundation's independent audits

There is a plethora of social auditing initiatives in the apparel industry, ranging from corporate-driven industry initiatives to multi-stakeholder initiatives (Anner, 2012; O'Rourke, 2006). Scholars discussing the impact of these initiatives stress the importance of a strong role for progressive NGOs and labor unions in the initiatives (Anner, 2012; O'Rourke, 2006), and they criticize corporate-influenced initiatives such as FLA, Business Social Compliance Initiative (BSCI), and Global Social Compliance Programme (GSCP) for their limited impact on worker rights (Anner, 2012; Egels-Zandén and Merk, 2013). Anner (2012), for example, attributes the failure of FLA audits to identify and remediate freedom of association violations to the significant role corporations play in the design and oversight of FLA.

Fair Wear Foundation is a Dutch-based European non-profit multi-stakeholder initiative that conducts independent verification of European companies' performance and efforts in securing worker rights in line with the FWF code of conduct. FWF was established in 1999 by the Dutch Clean Clothes Campaign in collaboration with labor union representatives and retailer associations; it has a strong role for labor organizations in both the design and oversight of the organization (O'Rourke, 2006). For example, FWF's board and ownership are divided equally between corporate organizations (only industry associations) and labor organizations (labor unions and human rights NGOs) (FWF, 2009). FWF also limits undue influence from individual companies by obtaining funding from different stakeholder groups (FWF, 2011). FWF is thus substantially different from corporate-influenced initiatives such as

FLA and BSCI. The nature of FWF and its audit methodology also makes its audits substantially different from the corporate-driven audits used in previous research (Locke et al., 2007a, 2007b).

FWF member companies are required to implement the FWF code of conduct at their suppliers. The FWF code comprises the relevant ILO conventions and covers eight areas:

1. *Forced labor*: There shall be no use of forced, including bonded or prison, labor (ILO Conventions 29 and 105).
2. *Child labor*: The age for admission to employment shall not be less than the age of completion of compulsory schooling and, in any case, not less than 15 years (ILO Convention 138 and 182).
3. *Discrimination*: Practices shall be based on the principle of equal opportunities (ILO Conventions 100 and 111).
4. *Trade union rights and worker committees*: The right of all workers to form and join trade unions and bargain collectively shall be recognized (ILO Conventions 87 and 98). In situations with legal restrictions, companies shall facilitate parallel means of independent and free association and bargaining for all workers (ILO Convention 135 and Recommendation 143).
5. *Wages*: Wages and benefits paid for a standard working week shall meet at least legal or industry minimum standards and always be sufficient to meet basic needs of workers and their families and to provide some discretionary income (ILO Conventions 26 and 131).
6. *Working times*: Hours of work shall comply with applicable laws and industry standards. In any event, workers shall not on a regular basis be required to work in excess of 48 hours per week and shall be provided with at least one day off for every seven-day period. Overtime shall be voluntary, shall not exceed 12 hours per week, shall not be demanded on a regular basis and shall always be compensated at a premium rate (ILO Convention 1).
7. *Health and safety*: A safe and hygienic working environment shall be provided (following ILO Convention 155 among other standards).
8. *Legally binding employment relationships*: Obligations to employees under labor or social security laws and regulations shall be respected and there shall be no excessive use of contract workers or apprenticeship schemes.

FWF member companies are required to include the FWF code in their supplier contracts, ensure that the code is posted at supplier factories, audit suppliers, and undertake corrective actions, i.e., to work with supply chain sustainability in ways similar to most other brands in the garment and apparel industry (Mamic, 2005). FWF then implements factory audits to control the quality of the member companies' own auditing, to follow up remediation after a complaint has been filed by a worker at a factory, or to verify working conditions at a randomly chosen factory (FWF, 2010). The member company can also ask FWF to do additional audits on top of the audits done by FWF for verification purposes. These audits are then paid for by the member company.

FWF's audit methodology was developed in 2004 (FWF, 2005). All factory audits are done by three local experts who each are specialized in one of the following areas: worker interviews, health and safety, or document inspection. The audits are preannounced, but, in contrast to most corporate-driven audits, worker off-site interviews without the involvement of factory management are conducted prior to the factory audit. Interviews with local stakeholders are also conducted as complements to worker interviews. During the audits, information is gathered from inspection of production facilities, interviews with management and workers, and document inspection. Conclusions are drawn regarding legal and FWF code compliance based on these five sources of information (stakeholders, workers, management, documents, and inspection of facilities). On average, nine and a half person-days are used for each factory audit (FWF, 2005). These measures lead to FWF having among the most, if not the most, stringent and credible factory audits in the garment industry.

The fact that FWF's audits are among the most stringent and credible factory audits does not dismiss the critique that factory audits, regardless of their stringency, are unable to capture working conditions. In other words, even well-trained, well-resourced, and dedicated auditors —with a clear commitment from their clients to assess compliance—may face serious difficulties in assessing supplier compliance. Numerous authors, as discussed in this paper, have stressed the limits of factory auditing (e.g., Anner, 2012; Brown, 2013; Egels-Zandén, 2007) even to the extent that the merit of the code of conduct compliance model has been questioned (Locke et al., 2009). However, given the extensive time it takes to unveil a supplier's compliance level, it is near impossible for even the most skilled team of researchers to themselves collect data about a large number of suppliers (thus, the focus on small number of cases in previous research). Given this situation, quantitative studies in the near future are

likely to continue to rely on secondary data. The strength of the FWF's audit data used in this paper is that it represents among the most credible of such secondary data.

4. Method

To advance research into the impact of codes of conduct requires access to multiple reliable, detailed, and traceable audits of supplier factories. These requirements were met in the proprietary dataset we obtained from FWF, which consisted of 288 full audit reports from audits conducted at 229 garment factories between 2004 and 2012, i.e., all audits conducted by FWF between 2004 and 2012. While all 288 audit reports were coded, the focus of this paper is the 43 factories that were audited multiple times by FWF (on average 26 months between the first and second audit). The audits for these 43 factories were coded twice by different persons and the results compared to derive the final version used in the statistical analysis in this paper. The reason for this limited focus is that only for these 43 factories did we have two reliable measuring points obtained through multiple audits using a consistent methodology.

The majority of the 43 factories studied were situated in China (42%), Tunisia (12%), Turkey (12%), Macedonia (12%), and India (7%). The rest of the factories were situated in Bangladesh, Bulgaria, Moldavia, Poland, Romania, and Thailand. The audited factories had an average of 298 workers (n = 43, min = 17, max = 2260, median = 189). The factories had on average been in operation for 10 years (n = 43, min = 1 max = 41, median = 10) and had on average been cooperating with the FWF member company for 4.5 years (n = 13, min = 0, max = 10, median = 3). Of the factories studied, 86% were direct suppliers to the FWF member, while 14% were tier-two suppliers; 7% had ISO9001 certifications and 5% had SA8000 certifications.

The audit reports were coded based on factory characteristics (geographic location, number of employees, years of operation, supplier tier, certifications, and length of buyer–supplier relations), and the FWF code's eight areas (forced labor, child labor, discrimination, freedom of association, wages, working times, health and safety, and legally binding employment relationships). In FWF's audit reports, each of the eight code areas were made up of between two (forced labor and discrimination) and 17 (health and safety) sub-categories (in total, 53 sub-categories for the eight areas). As FWF audit protocols up to 2009 listed remarks on

committees and grievance mechanisms under freedom of association (FoA), this classification was used for all subsequent audits as well; we refer to this area as "freedom of association and committees."

To assess the extent to which factories complied with codes of conduct, we constructed four different measurements. In line with Oka (2010a, 2010b) and Toffel et al. (2012), we calculated the total number of violations identified in the factory audit, i.e., the sum of all identified violations without any weighing of the importance of the violation. While the total number of violations provides a rough compliance estimate, it is a problematic measure for several reasons. Most notably, auditors tend to differ in how many minor non-compliances they include in the audit report, and they tend to group these minor non-compliances differently (some listing them individually and some grouping them in a handful of categories). This makes the total number of identified violations a relatively unreliable compliance measurement.

We, therefore, developed a second compliance measurement by assigning each factory a score of 0–3 for each of the eight areas based on the factory's compliance in the 53 sub-categories (0 = fully compliant and 3 = highly non-compliant). Appendix 1 outlines the scoring criteria. This scoring allowed us to discuss both compliance in each of the eight FWF code areas and total compliance (measured as the combined score for the eight areas, i.e., 8*0–3 = 0–24 points). In other words, it allowed us to test both the argument that codes have uneven impact and the overall impact of codes. Finally, since auditors more consistently noted major non-compliances (score points 2–3 in our measurement) as compared to minor non-compliances (score 1 in our measurement), we developed a measure that solely included major non-compliances (i.e., we used a 0–2 scale where minor non-compliances (score 1) were coded as fully compliant (score 0)). To examine the impact of codes of conduct, we analyzed whether there was a difference in the 43 suppliers' compliance (as measured by the three developed measurements) between the first and second FWF audit.

In terms of internal validity, our empirical data were limited in two important ways. First, and in contrast to, for example, Oka (2010a, 2010b), we lacked a control group of suppliers not exposed to codes of conduct, i.e., we had difficulty controlling for other variables influencing supplier compliance over time. Second, we lacked data on working conditions prior to the first conducted FWF audit. Since some buyers signaled that they demanded compliance

before placement of orders, it is possible that codes could have had an impact prior to the first audit. Oka (2010a, 2010b), for example, argues that such pre-order and pre-audit effects improve compliance. Our data were limited to analyzing post-order and post-audit improvements. The implications of these limitations are further discussed in the final sections of this paper.

In terms of external validity, FWF member companies and suppliers work with sustainable supply chains in similar ways to that of most other garment and apparel companies (Mamic, 2005; Starmanns, 2010). Still, as noted above, FWF differs from some of the other multi-stakeholder initiatives in the garment industry such as BSCI and FLA. It is thus reasonable to assume that a different type of company joins FWF as compared to, for example, BSCI. Given that FWF, for example, demands more stringent factory auditing than most other multi-stakeholder initiatives, it could be characterized as a favorable setting for examining the impact of codes. The strength of this "critical case" approach is that *if* it is shown that codes in the favorable setting do *not* improve compliance, it is reasonable to assume that codes in less favorable settings (such as BSCI and FLA member companies) will not improve compliance either (Flyvbjerg, 2001). Hence, this approach provides a way to generalize through logical deduction based on a rationale closely linked to Karl Popper' s notion of falsification (Flyvbjerg, 2001). Of course, this also implies that the approach does not allow for generalizations if codes in the favorable setting improve compliance, since this provides little information regarding improved compliance in less favorable settings. The ability of our findings to speak more generally to the garment and apparel industry, thus, depends on the results themselves. Regardless of results, our study is, of course, also limited by our exclusive focus on the garment industry since industry differences have been shown to influence, at least, suppliers' code of conduct adherence (Toffel et al., 2012).

5. Results

5.1. Identified non-compliances

Given the quality problems of factory audits identified in previous research, we assessed the quality of the FWF audits by analyzing the findings from all 229 factories in all 60 sub-categories of the eight code areas. The results were striking in that three sub-categories showed few violations: discrimination of trade union members or management interference in trade union (found at 4% of factories), harassment (found at 2% of factories), and gender

discrimination (found at 12% of factories). Since FWF's audit reports include interviews with relevant local stakeholders, these findings could be compared with expressed stakeholder views. Violations to freedom of association, gender discrimination, and harassment were all listed by the interviewed stakeholders as high risk areas of non-compliance in the garment industry in the countries studied. Hence, there was a discrepancy between what local stakeholders expected FWF auditors to find and what was actually found by the auditors. Interestingly, in other high risk areas of non-compliance identified by the local stakeholders (such as health and safety, working times, and wages), audits mirrored stakeholder expectations, with no factory (out of 229) being fully compliant in terms of health and safety, 88% of the factories receiving remarks on working times, and 81% receiving remarks on wages.

In sum, while FWF's audits are generally regarded as among the most rigorous and credible factory audits, the audits identified surprisingly few non-compliances in freedom of association, harassment, or gender discrimination. This lack is potentially problematic, given local stakeholders' expectations and will be discussed in more detail at the end of this paper.

5.2. Improvements between first and second audit?

As a first test of whether or not supplier compliance improved between the first and second audit, we compared the total number of non-compliances at the first and second audit. A Wilcoxon signed ranks test showed significant improvements between the first and second audit: factories improved on average from 22.1 remarks to 18.8 remarks ($p = 0.007$, $N = 43$).

Since the total number of violations is a crude measurement of compliance, we also compared the total grade for each factory, based on the grading outlined in Appendix 1: i.e., min = 0 (fully compliant in all eight areas), max = 24 (fully non-compliant in all eight areas). Again, the Wilcoxon signed ranks test showed that the decrease from the first audit, mean grade 12.3, to the second audit, mean grade 10.6, was statistically significant ($p = 0.005$, $N = 43$).

Finally, we tested the overall impact of codes solely for major non-compliances (i.e., scoring minor non-compliances (1) as compliances (0), creating a grade scale from fully compliant = 0 to fully non-compliant = 16). Again, the Wilcoxon signed ranks test showed significant improvements between the first and second audit ($p = 0.010$). Hence, regardless of

compliance measure, our results indicate that factories undergoing code of conduct audits improve working conditions over time.

To test the potentially unequal impact of codes, differences between the first and second audit were also analyzed for each of the eight code areas, using (i) the total number of remarks, (ii) the 0–3 grading system (Appendix 1), and (iii) the 0–2 grading system focusing on major non-compliances. Table 2 summarizes the changes in mean number of total remarks between the first and the second audit. The average number of remarks decreases in all areas except for freedom of association and committees where there is a marginal increase. However, the only statistically significant change is the number of remarks on child labor which decreases from on average 0.93 to 0.56 ($p = 0.015$).

Insert Table 2 about here

Table 3 summarizes the findings for each of the eight areas based on the 0–3 grading. In all eight areas, except for working times, suppliers on average improved. For working times there was no change at all. The employment relationship area experienced the biggest improvement (mean first audit = 2.19, mean second audit = 1.79), and a Wilcoxon signed ranks test showed this was the only area with a statistically significant change between the first and second audits ($p = 0.036$).

Insert Table 3 about here

Table 4 summarizes the findings in terms of changes in serious violations for each of the eight areas. In all eight areas, except for working times, suppliers on average improved. For working times, there was a somewhat declined performance. Health and safety showed the biggest improvement (mean first audit = 0.95, mean second audit = 0.7), but a Wilcoxon signed ranks test showed no significant results for any of the eight areas.

A closer look at the areas where significant improvements were found, i.e., child labor (in terms of total number of remarks) and legally binding employment relationship (when comparing grades 0–3), reveal the reasons for the improvements. For child labor, every factory that received a remark on having children working had improved at the time of the second audit; for legally binding employment relationship, the major improvement was that workers received contracts to a greater extent at the second audit.

Insert Table 4 about here

To more fully understand what happens in terms of supplier compliance between the first and second audit, we analyzed how many factories showed improved, declined, or unchanged compliance between the audits. Table 5 summarizes the results: many of the suppliers that were non-compliant in a specific area improved their compliance in this area between the first and second audit (min = 21% improved for working times, max = 76% improved for discrimination). However, in parallel with these improvements, many of the compliant suppliers in a specific area became non-compliant suppliers between the first and second audit. In all cases, about 10% at least of the suppliers showed declined performance between the first and second audit.

Insert Table 5 about here

---

6. Discussion

6.1. Auditing more flawed than previously assumed

The results show that FWF factory audits only identify freedom of association (FoA) violations and harassment in less than five percent of the factories (in four and two percent respectively), and gender discrimination in around ten percent of the factories. This result is in line with Anner's (2012) finding that Fair Labor Association audits rarely identify FoA violations, and Egels-Zandén and Merk's (2013) argument that factory audits are unlikely to detect FoA violations. It is also in line with Barrientos and Smith's (2007) finding that

discrimination and harassment are unlikely to be addressed through codes of conduct. The lack of identified violations in FWF's factory audits is problematic, given the interviewed stakeholders' expectations of extensive violations of this type and that other studies have shown that such violations are common in the countries where the suppliers in the study were located (Anner, 2012). It is thus likely that FWF audits underestimate the number of FoA and discrimination violations at supplier factories. This, in turn, supports the widely-held view that codes of conduct are unable to improve process rights (Barrientos and Smith, 2007) since the ability to identify violations is a necessary requirement for potential improvements.

This result is distressing, since previous research has largely attributed factory audits' inability to detect FoA and discrimination violations to corporate-led auditing (Anner, 2012; Egels-Zandén and Merk, 2013). For example, Anner (2012) argues that the inability of audits to detect FoA violations is due to FLA being a corporate-influenced initiative, and that such initiatives can be expected to prioritize outcome standards as compared to process rights. Our results show that this is a too optimistic explanation and that factory audits are flawed on a more fundamental level. Progressive NGOs and labor unions have a strong role in FWF and thus may be expected to stress the importance of FoA and discrimination. Still, such violations are not identified, indicating that even the most rigorous factory audits (including off-site interviews with workers, etc.) conducted by initiatives with a strong labor rights agenda are insufficient for identifying and addressing FoA and discrimination in supply chains.

6.2. Codes have positive (although limited) overall impact

Our results of the overall impact of codes of conduct show that suppliers, regardless of measurements, improve between the first and second audit. This finding stands in contrast to Locke et al.'s (2007a, 2007b) finding that the compliance of Nike suppliers did not improve over time. There are several potential explanations for these different findings. One explanation is that the difference is due to time period differences: Locke et al. (2007a, 2007b) covered an early code of conduct phase (1998–2005) and the FWF study a more recent phase (2002–2012). In other words, the commonly voiced objection by corporate managers that Locke et al.'s findings are outdated and that codes and auditing in later years have had greater impact is potentially supported by our results. This interpretation is also in line with Egels-Zandén's (2013) finding that codes had greater impact in 2009 than in 2004 in the Chinese toy industry.

Another possible explanation is that Locke et al.'s findings were specific to Nike (a large U.S.-based company that has been one of the main targets of the anti-sweatshop movement) and that codes are more effective in the hands of the more proactive European companies that are members of FWF. Yet another possibility is that it is the type of auditing data that explains the difference, with Locke et al.'s (2007a, 2007b) study being based on corporate-driven audits and our study on multi-stakeholder-driven audits.

While we find statistically significant improvements between the first and second audit, the actual improvements over time are limited. On average, suppliers had 15% fewer total remarks in the second audit (mean first audit = 22.05, mean second audit = 18.77), 14% improved scores based on the 0–3 grading (mean first audit = 12.3, mean second audit = 10.6), and 7% improved scores based on the 0–2 grading (mean first audit = 2.99, mean second audit = 2.77). Given that 26 months on average had passed between the first and second audit, our findings still raise doubts about the effectiveness of codes and auditing. In other words, despite finding statistically significant improvements, our results partly support the claim that codes of conduct, at least post-order and post-audit, are too weak for the job of substantially improving worker rights in global production networks (Blowfield and Dolan 2008; Chan and Siu, 2010; Frundt, 2004).

6.3. General but not specific uneven impact

Our findings both support and challenge the argument that codes improve outcome standards but not process rights (Barrientos and Smith, 2007). The findings provide support in the sense that, as discussed here, factory audits seem unable to even identify process rights violations. Furthermore, although less relevant given the auditing problems, we find no statistically significant improvements in process rights such as freedom of association or discrimination. The present findings also support the finding in previous case study research that codes improve outcome standards (Barrientos and Smith, 2007; Egels-Zandén, 2013; Frenkel, 2001; Mamic 2004), since we identify significant improvements in terms of outcome standards in aggregate. However, in relation to specific outcome standards, we only find statistically significant improvements in two (child labor and employment relationships) of the six outcome standards we examined (forced labor, child labor, wages, working times, health and safety, and employment relationships). Additionally, in these two outcome standards, we only

find significant results in one of the three developed measures (total number of remarks for child labor, and 0–3 grading for employment relationships).

Our findings, thus, indicate that codes can improve outcome standards in aggregate, but that such overall improvements are limited (see section "Codes have positive (although limited) overall impact") and that few statistically significant improvements are observed in relation to specific outcome standards. Given that our study is the first quantitative effort to test the argument of uneven impact and that the results are not clear-cut, further studies in both the garment and other industries are needed to shed light on this debate.

6.4. The illusion of improvement

In addition to partly challenging the ability of codes to address specific outcome standards, our results also challenge companies' multi-billion investments in codes and auditing. If codes only have limited overall impact, why do companies (and multi-stakeholder initiatives) continue to invest in codes and auditing? Are corporate managers knowingly investing in codes and auditing with limited impact simply to minimize brand risk and avoid activist campaigns? Although this is a possibility, a less cynical interpretation is that sustainability managers and representatives of multi-stakeholder initiatives actually perceive that codes and auditing do make a difference at the point of production.

Our results present a possible explanation for the conflicting views about the impact of codes. As shown in Table 5, many of the suppliers identified as non-compliant in an area during the first audit showed improved performance in that area in the second audit. In other words, if we simply compare non-compliant supplier performance in the areas of non-compliant between the first and second audit, codes and auditing seem to have an impact. However, in parallel to these improvements, compliant suppliers experience declined performance, reducing the overall improvements. A single supplier, thus, shows both improved and declined performance over time.

An auditor, sustainability manager, or multi-stakeholder representative conducting or analyzing follow-up factory audits is likely to specifically note changes in the previously identified non-compliance areas and is then likely to interpret the improvement "cues" (Weick, 1995) in these areas as signs of codes and auditing improving working conditions. The cues of declined performance are likely to either be disregarded or interpreted as signs

that further auditing is needed (since this seems to have worked for previous non-compliant areas). In this way, an illusion of improvement is created amongst those that work with codes of conduct and auditing, while actual worker rights do not improve to the same extent over time.

The finding that suppliers move in and out of compliance shows that the main problem with codes and auditing is not that they are incapable of addressing non-compliance, since our results show that they actually do exactly this. Instead, the main problem is that codes and auditing are unable to sustain compliance, i.e., unable to make a compliant supplier stay compliant. One interpretation of this finding is that codes and auditing are unable to hinder contradictory tendencies in the garment industry such as flexibilization of labor markets, declining purchasing prices, and shorter lead times (Bartley and Kincaid, 2013). In other words, as these tendencies increase, they undermine the effects of codes and audits by turning compliant suppliers into non-compliant suppliers (Bartley and Kincaid, 2013). This opens the door to the positive interpretation that codes and auditing potentially either mitigate this performance decline and/or improve areas not affected by contradictory tendencies. Such areas could include improved policies and documentation that could be decoupled from substantive actions in the operative core of the supplier's operations (Meyer and Rowan, 1977). If so, suppliers experiencing codes and auditing should perform better than those not experiencing codes, even if those experiencing codes only show limited improvements over time.

Another interpretation of why codes are perceived as improving worker rights while this and other studies find no or limited improvements is that codes mainly have a pre-first-audit impact on worker rights. In other words, managers correctly perceive that suppliers improve, but this improvement is made between the initial informal contact with the supplier and the first formal factory audit. Both these interpretations reconcile the contradictory finding in previous research that suppliers exposed to codes show better worker rights compliance (Oka, 2010a, 2010b), even though codes either fail to improve worker rights over time (Locke et al., 2007a, 2007b) or only improve worker rights to a limited extent, as shown in our FWF study. Further research could test the validity of these interpretations by, for example, comparing the sustainability performance of audited and non-audited factories over time and preferably including a pre-first-audit measurement of sustainability performance.

7. Conclusion

The rise of private regulation of sustainability in global production networks has led to intensive debates about the impact of this regulation at the point of production. Few empirical studies have moved beyond qualitative studies of a small number of firms and assessed the impact of private regulation of worker rights in practice. Drawing on multiple factory audits of 43 garment factories, we show that codes improve worker rights on an overall level, although marginally. Our findings also both support and partly challenge the claim in previous research (Barrientos and Smith, 2007) that codes have uneven impact. The findings support for our position that even rigorous multi-stakeholder factory audits are unable to identify process rights violations (such as FoA or discrimination violations) and that auditing is thus more fundamentally flawed than assumed in previous research (Anner, 2012). The findings also support the argument of statistically significant improvement in terms of outcome standards in the aggregate, but offer a challenge in the sense that we hardly find any statistically significant improvement in specific code areas. Given the limited overall improvements that codes lead to, it is surprising that companies continue their multi-billions of investments in codes and auditing. Our findings indicate that this contradiction is explained by either an illusion of improvement, with actors focusing on corrected non-compliances rather than the new non-compliances that emerge in parallel, and/or pre-first-audit improvements and/or that codes mitigate overall compliance decline.

These results and the limitations of our used method suggest several avenues for future research. Given that both the present study and Locke et al.'s (2007a, 2007b) focus on the apparel industry, quantitative impact studies in other industries are needed to determine whether codes have more substantial effects in other industries. For example, it is reasonable to assume that codes will have more limited impact in industries where sustainability demands are less forceful than in the garment industry (cf. Flyvbjerg, 2001), while potentially having more impact in industries with less of a conflict between sustainability and purchasing criteria (cf. Bartley and Kincaid, 2013) and with suppliers located in countries with more stringent employment law regimes, high press freedom and many international NGOs (cf. Toffel et al., 2012).

Furthermore, our limited sample does not allow us to examine whether codes of conduct could improve worker rights under specific circumstances (such as specific geographic

locations, factory sizes, and types of buyer–supplier relations). The emerging literature into when suppliers comply with codes of conduct (e.g., Toffel et al., 2012) could though address this limitation if expanded to include improvements over time as well. Furthermore, our method neither allows for tests of pre-first-audit effects by comparing pre-audit compliance with post-audit compliance nor for analysis of whether or not codes mitigate compliance decline by comparing suppliers exposed, and not exposed, to codes over time. Future research should include such measures to more fully capture the impact of codes. Further studies are also required into the potentially uneven impact of codes of conduct. Do our findings of non-significant improvements in specific worker rights areas hold even in other circumstances (such as other industries, geographic locations, or other forms of private regulation such as certifications)?

For managers, the findings are both positive and distressing, since they present a somewhat more positive picture than Locke et al. (2007a, 2007b), but still question the ability of even the most rigorous codes and factory auditing to, at least in the garment industry, substantially improve worker rights in global production networks. It is therefore likely that the less rigorous codes and audits that most companies, in the garment and other related consumer industries, invest in have similarly limited impact (cf. Flyvbjerg, 2001). Managers have to decide if they should move beyond the compliance model of codes into, for example, more cooperative models (Locke et al., 2009), although such cooperative models also face criticism (Lund-Thomsen and Lindgreen, 2013); whether they should abandon codes for other forms of private regulation (such as certifications); or whether there are ways to make codes effective. Regardless of the chosen path forward, managers are well advised not to accept the status quo since codes and factory auditing in the present form have difficulties in substantially improving worker rights.

8. References

Anner, M., 2012. Corporate social responsibility and freedom of association rights : The precarious quest for legitimacy and control in global supply chains. Politics & Society 40(4), 609-644.

Armbruster-Sandoval, R., 2005. Globalization and Cross-border Labor Solidarity in the Americas: The Anti-sweatshop Movement and the Struggle for Social Justice. Routledge, New York.

Auret, D., Barrientos, S., 2004. Participatory social auditing: A practical guide to developing a gender-sensitive approach. IDS Working Paper 237, Institute of Development Studies, Brighton.

Barrientos, S., Gereffi, G., Rossi, A., 2011a. Economic and social upgrading in global production networks: A new paradigm for a changing world. International Labour Review 150(3–4), 319-340.

Barrientos, S., Mayer, F., Pickles, J., Posthuma, A., 2011b. Decent work in global production networks: Framing the policy debate. International Labour Review 150(3–4), 299-317.

Barrientos, S., Smith, S., 2007. Do workers benefit from ethical trade? Assessing codes of labour practice in global production systems. Third World Quarterly 28(4), 713-729.

Bartley, T., 2007. Institutional emergence in an era of globalization: The rise of transnational private regulation of labor and environmental conditions. American Journal of Sociology 113(2), 297-351.

Bartley, T., Kincaid, D., 2013. The mobility of industries and the limits of CSR: Labor codes of conduct in Indonesian factories. Unpublished Working Paper.

Blowfield, M., Dolan, C., 2008. Stewards of virtue? The ethical dilemma of CSR in African horticulture. Development and Change 39(1), 1–23.

Boring, E.G., 1954. The nature and history of experimental control. American Journal of Psychology 67(4), 573-589.

Brown, G., 2013. Fatal flaws of foreign factory audits. Industrial Safety and Hygiene News, February. Accessed at: http://www.ishn.com/articles/print/95045-fatal-flaws-of-foreign-factory-audits

Chan, A., Siu, K., 2010. Analyzing exploitation: The mechanisms underpinning low wages and excessive overtime in Chinese export factories. Critical Asian Studies 42(2), 167-190.

Egels-Zandén, N., 2007. Suppliers' compliance with MNCs' codes of conduct: Behind the scenes at Chinese toy suppliers. Journal of Business Ethics 75(1), 45-62.

Egels-Zandén, N., 2013. Revisiting supplier compliance with MNC codes of conduct: Recoupling policy and practice at Chinese toy suppliers. Journal of Business Ethics (in press).

Egels-Zandén, N., Merk, J., 2013. Private regulation and trade union rights: Why codes of conduct have limited impact on trade union rights. Journal of Business Ethics (in press).

Frenkel, S., 2001. Globalization, athletic footwear commodity chains and employment relations in China. Organization Studies 22(4), 531-562.

Frundt, H.J., 2004. Unions wrestle with corporate codes of conduct. WorkingUSA 7(4), 36-69.

FWF (Fair Wear Foundation)., 2005. Fair wear foundation audit manual. Available at http://www.fairwear.org/ul/cms/fck-uploaded/documents/policydocs/FWFauditmanual-march2012.pdf. Accessed on May 16, 2013.

FWF (Fair Wear Foundation)., 2009. Fair wear foundation: Charter. Available at http://www.fairwear.org/ul/cms/fck-uploaded/documents/companies/FWFdocs/fwfcharterjanuary2011.pdf. Accessed on May 16, 2013.

FWF (Fair Wear Foundation)., 2010. The fair wear formula. Available at http://www.fairwear.org/ul/cms/fck-uploaded/documents/fwfpublications_reports/thefairwearformula.pdf. Accessed on May 16, 2013.

FWF (Fair Wear Foundation)., 2011. Fair wear foundation: Annual report 2010. Available at http://www.fairwear.org/ul/cms/fck-uploaded/documents/fwfpublications_reports/FWFannualreport2010.pdf. Accessed on May 16, 2013.

Jiang, B., 2009. Implementing supplier codes of conduct in global supply chains: Process explanations from theoretic and empirical perspectives. Journal of Business Ethics 85, 77-92.

Locke, R.M., Amengual, M., Mangla, A., 2009. Virtue out of necessity? Compliance, commitment and the improvement of labor conditions in global supply chains. Politics & Society 37(3), 319-351.

Locke, R.M., Kochan, T., Romis, M., Qin, F., 2007a. Beyond corporate codes of conduct: Work organization and labour standards at Nike's suppliers. International Labour Review 146(1–2), 21-40.

Locke, R.M., Qin, F., Brause, A., 2007b. Does monitoring improve labor standards? Lessons from Nike. Industrial and Labor Relations Review 61(1), 3-31.

Locke, R.M., Rissing, B.A., Pal, T., 2013. Complements or substitutes? Private codes, state regulation and the enforcement of labour standards in global supply chains. British Journal of Industrial Relations (in press).

Lund-Thomsen, P., Lindgreen, A., 2013. Corporate social responsibility in global value chains: Where are we now and where are we going? Journal of Business Ethics (in press).

Mamic, I., 2004. Implementing codes of conduct: How businesses manage social performance in global supply chains. Greenleaf Publishing, Sheffield, UK.

Mamic, I., 2005. Managing global supply chain: The sports footwear, apparel and retail sectors. Journal of Business Ethics 59, 81-100.

Oka, C., 2010a. Accounting for the gaps in labour standard compliance: The role of reputation-conscious buyers in the Cambodian garment industry. European Journal of Development Research 22(1), 59–78.

Oka, C., 2010b. Channels of buyer influence and labor standard compliance: The case of Cambodia's garment sector. Advances in Industrial and Labor Relations 17,153–183.

O'Rourke, D., 1997. Smoke from a Hired Gun: A Critique of Nike's Labor and Environmental Auditing in Vietnam as Performed by Ernst & Young. Transnational Resource and Action Center, San Francisco, CA.

O'Rourke, D., 2006. Multi-stakeholder regulation: Privatizing or socializing global labor standards? World Development 34(5), 899–918.

Pearson, R., Seyfang, G. 2001. New Hope or False Dawn? Global Social Policy 1(1), 49–78.

Rodríguez-Garavito, C.A., 2005. Global governance and labor rights: Codes of conduct and anti-sweatshop struggles in global apparel factories in Mexico and Guatemala. Politics and Society 33(2), 203-233.

Ross, R.J.S., 2006. A tale of two factories: Successful resistance to sweatshops and the limits of firefighting. Labor Studies Journal 30(4), 65-85.

Rowley, T., Berman, S., 2000. A brand new brand of corporate social performance. Business & Society 39(4), 397-418.

Ruggie, J., 2004. How to marry civic politics and global governance, in Kuper, A. (Ed.), The impact of corporations on global governance. Carnegie Council on Ethics and International Affairs, New York, pp. 10–23.

Seuring, S., Gold, S., 2013. Sustainability management beyond corporate boundaries: From stakeholders to performance. Journal of Cleaner Production (in press).

Seuring, S., Müller, M., 2008. From a literature review to a conceptual framework for sustainable supply chain management. Journal of Cleaner Production 16, 1699-1710.

Seuring, S., Sarkis, J., Müller, M., Rao, P., 2008. Sustainability and supply chain management: An introduction to the special issue. Journal of Cleaner Production 16, 1545-1551.

Stouffer, S.A., 1949. The American Soldier. Princeton University Press, Princeton.

Taylor, M., 2011. Race you to the bottom … and back again? The uneven development of labour codes of conduct. New Political Economy 16(4), 445-462.

Toffel, M.W., Short, J.L., Ouellet, M., 2012. Reinforcing regulatory regimes: How states, civil society, and codes of conduct promote adherence to global labor standards. Working Paper 13-045. Harvard Business School.

Weick, K.E., 1995. Sensemaking in organizations. Sage, Thousand Oaks, CA.

Wells, D., 2007. Too weak for the job: Corporate codes of conduct, non governmental organizations and the regulation of international labour standards. Global Social Policy 7(1), 51–74.

Zadek, S., 2004. The path to Corporate Social Responsibility. Harvard Business Review 82(12), 125–132.

Tables

Table 1. Main differences between previous studies and this study

	Previous studies	**This study**
Type of study	Qualitative case studies	Quantitative
Type of data collection	Single point in time	Two points in time
Type of audits	Commercial actors	Multistakeholder initiative
Type of measures	Aggregated scores	Specific issue scores
Type of data	Aggregated suppliers	Specific suppliers

Table 2. Change in number of remarks between first and second audit for each labor standard (N=43)

	Mean number of remarks			
Labor standard	*First audit*	*Second audit*	*Change*	*P*
Forced labor	0.23	0.21	0.02	0.763
Discrimination	0.40	0.23	0.17	0.149
Child labor	0.93	0.56	0.37	0.015
FoA & committees	0.05	0.07	-0.02	0.317
Wages	2.81	2.79	0.02	0.967
Working times	3.70	3.49	0.21	0.554
Health & safety	8.16	7.16	1	0.99
Employment relationship	1.91	1.51	0.40	0.18
Mean total grade	22.05	18.77	3.28	0.007

Table 3. Change in grade between first and second audit for each labor standard using grading scale 0–3 (N=43)

Labor standard	*Grade*	*First audit grade*	*Second audit grade*	*Change*	*P*[1]
Forced labor	Mean	0.33	0.28	0.05	0.747
	0	79.1	76.7		
	1	11.6	18.6		
	2	7	4.7		
	3	2.3	0		
Discrimination	Mean	0.81	0.42	0.39	0.069
	0	51.2	76.7		
	1	11.6	11.6		
	2	27.9	4.7		
	3	9.3	7		
Child labor	Mean	0.95	0.72	0.23	0.186
	0	51.2	60.5		
	1	11.6	7		
	2	27.9	32.6		
	3	9.3	0		
FoA & committees	Mean	1.65	1.49	0.16	0.216
	0	7	14		
	1	30.2	32.6		
	2	53.5	44.2		
	3	9.3	9.3		
Wages	Mean	2.23	2	0.23	0.180
	0	9.3	20.9		
	1	7	2.3		
	2	34.9	32.6		
	3	48.8	44.2		
Working times	Mean	2.19	2.19	0	0.963
	0	11.6	16.3		
	1	18.6	11.6		
	2	9.3	9.3		
	3	60.5	62.8		
Health & safety	Mean	1.95	1.7	0.25	0.065
	0	0	0		
	1	30.2	44.2		
	2	44.2	41.9		
	3	25.6	14		
Employment relationship	Mean	2.19	1.79	0.4	0.036
	0	9.3	25.6		
	1	7	4.7		
	2	39.5	34.9		

[1] Wilcoxon Signed Ranks Test: symptotic significance, 2-tailed

	3	44.2	34.9		
Mean total grade		12.3	10.6	1.7	0.005

Table 4. Change in serious violations between first and second audit for each labor standard using grading scale 0–2 (N=43)

Labor standard	*First audit grade*	*Second audit grade*	*Change*	*P*
Forced labor	0.12	0.05	0.07	0.317
Discrimination	0.42	0.19	0.23	0.097
Child labor	0.47	0.33	0.14	0.109
FoA & committees	0.72	0.63	0.09	0.346
Wages	1.33	1.21	0.12	0.251
Working times	1.3	1.35	-0.05	0.557
Health & safety	0.95	0.7	0.25	0.065
Employment relationship	1.28	1.05	0.23	0.082
Mean total grade	2.99	2.77	0.22	0.01

Table 5. Compliance changes between first and second audit (N=43)

Percentage of factories where grades:	*Forced labor*	*Discrimination*	*Child labor*	*FoA & commitees*	*Wage*	*Working time*	*H&S*	*Legal employment*
Improved	14%	30%	23%	26%	21%	19%	26%	35%
Declined	14%	9%	9%	14%	16%	12%	19%	12%
Did not change: no remarks	65%	49%	42%	5%	2%	9%	0%	9%
Did not change: non-compliant factory	7%	12%	26%	56%	60%	60%	56%	44%
Percentage of factories with remarks at the first audit that had improved at the second audit	67%	76%	48%	30%	23%	21%	26%	38%

Appendix 1

Explanation of 0–3 grading system

Non-compliances	Grades / non-compliance
No forced labor	
No remark	0
Workers not allowed to leave until finished quota	1
No policies on forced labor	1
Too long resignation time	1
Restrictions on when holidays can be taken	1
Deposits taken for tools used	1
Payments done late in order to make workers stay	2
Salaries withheld	2
Withholding salary or legal benefits if a worker quits	2
Workers must get permission to quit	2
Workers must find replacement to quit	2
Workers are fired if they refuse to work OT	2
Workers are fined if refuse to do OT	2
Restrictions on leaving factory compound	2
Forced labor, i.e. locked doors, confiscated IDs, etc.	3
Workers need to pay if they quit	3
No discrimination	
No remark	0
Unclear payment policy, risk of discrimination	1
Workers have not understood policies on non-discrimination	1
Missing policies on non-discrimination	1
Men and women given different types of jobs	2
Different pay for same work based on kinship/friendship	2
Beliefs concerning gender restrict the ability of women to acquire skills and move into higher-paid positions	2
Facilities different for men / women	2
Workers (or those applying for jobs) have to do pregnancy tests	3
Pregnant workers are fired	3
Different payment for same job based on gender / ethnicity / etc	3
Only male supervisor in workplace with both men and women	3
Maternity benefits not provided	3
Age discrimination, for example not employing workers over certain age	3
Child labor	
No remark	0
Remarks on system to avoid child labor	1
Child labor (historic)	2
Non-compliance with regulation for juvenile work	2
Present child labor	3

Freedom of association and communication	
No remark	0
Remark on committees *or* complaint system	1
Remarks on committees *and* complaint system	2
Discrimination of trade union members	3
Management appoint trade union representatives	3
Wage 0–3	
No remarks	0
Workers do not receive correct, or any, payslips	1
Workers do not understand wages paid	1
Wages do not meet local stakeholders' estimate of a living wage	1
Not possible to assess wages due to non-transparent	2
Correct overtime not paid	2
Wages not paid on time	2
Wage deductions made	2
Paid holiday not provided according to law	2
Minimum wages for normal working hours not paid	3
Working times 0–3	
No remarks	0
No proper timekeeping system in the factory	1
Inconsistency was found between records	1
Not possible to assess working hours	1
Excessive OT	2
Working weeks of above 60 hours	2
Working weeks of above 70 hours	3
No rest day for every 7 days	3
Safe and healthy working environment (each non-compliance graded 1 for one non-compliance and 2 if two or more non-compliances found)	
No remarks	0
1–5 remarks	1
6–10 remarks	2
More than 10 remarks	3
Legally binding employment relationship	
No remarks	0
Incorrect personnel records	1
Overtime work consent not according to law	1
Not all legally required insurances paid	2
Contracts do not confirm with legal requirements (or appointment letters where these required), include too long probation period	2
Short-term contracts used	2
Social security not paid and registered correctly	2
Not all workers provided with contract (or appointment letter where required) *or* factory keeping signed termination of employment form *or* factory does not have all contracts	3
Workers do not receive a copy of their contract (or appointment letter where required)	3