Monitoring the Monitors:
How Social Factors Influence Supply Chain Auditors [†]

Jodi L. Short
UC Hastings College of the Law

Michael W. Toffel
Harvard Business School

Andrea R. Hugill
Harvard Business School

Supply chain auditors provide companies with key strategic information about the practices of their suppliers, yet little is known about what influences their ability to identify and report dangerous, illegal, and unethical behavior at factories. Drawing on insights from the literatures on street-level bureaucracy and on regulatory and audit design, we theorize and investigate the factors that shape the practices of private supply chain auditors. We find evidence that their reporting practices are shaped by an array of social factors, including an auditor's experience, gender, and professional training; ongoing relationships between auditors and audited factories; and gender diversity on audit teams. By providing the first comprehensive and systematic findings on supply chain auditing practices, our study suggests strategies for designing more credible monitoring regimes.

Keywords: industry self-regulation, auditing, codes of conduct, supply chains, corporate social responsibility, globalization

Multinational companies (MNCs) are increasingly being held accountable for the social and environmental practices in their supply chains by consumers, investors, activists, and governments. Companies that seek to market to socially conscious consumers or attract socially responsible investors must demonstrate that their products are ethically and sustainably produced. Companies that seek to insulate their valuable brands from the negative publicity stemming from "name and shame" campaigns (Bansal, 2005) have acceded to activists' demands to monitor and attempt to improve practices in their supply chains (Berrone *et al.*, 2013). Companies must increasingly comply with domestic laws that require due diligence and/or

[†] We gratefully acknowledge the research assistance of Melissa Ouellet, as well as that of Chris Allen, John Galvin, Erika McCaffrey, and Christine Rivera. Xiang Ao, Max Bazerman, Morris Ratner, Bill Simpson, and participants in the UC Hastings College of the Law Junior Faculty Workshop offered helpful comments. Harvard Business School's Division of Research and Faculty Development provided financial support.

disclosure of their supply chain practices (Zandvliet, 2011).[1]

In this business environment, suppliers' social practices carry both opportunities and risks for their multinational buyers. The benefits of socially responsible production can be substantial. In 2008, Fairtrade-certified products accounted for nearly $5 billion worth of sales worldwide. A growing number of institutional investors rely on environment and social governance metrics (ESG) to make investment decisions (US SIF, 2012). Conversely, suppliers' poor social performance carries significant downside risks, as multinational buyers are often held accountable by NGOs, trade associations, and the press (e.g., Galland, 2010; Duhigg and Barboza, 2012; AFL-CIO, 2013). In December 2013, four global retailers who sourced from suppliers in the Rana Plaza factory in Bangladesh agreed to contribute nearly $40 million to compensate families of the 1,100 workers who died when the factory collapsed earlier that year (Greenhouse, 2013).

To exploit these opportunities and manage these risks, many companies have adopted private, voluntary standards to govern production practices in their global supply chains. These include meta-standards that specify procedures for systematic approaches to environmental management (ISO 14001) and socially responsible management (ISO 28000); technical production standards promulgated by certification and labeling regimes such as the Forest Stewardship Council; and corporate supplier codes of conduct governing workplace and environmental conditions that MNCs contractually impose on their suppliers. MNCs that adopt standards and codes like these also typically employ independent third-party auditors to monitor their suppliers' compliance. Third-party certification schemes are intended to provide reliable

[1] For example, the U.S. Securities and Exchange Commission's Conflict Minerals Rule requires companies to disclose their use of minerals originating in the Democratic Republic of Congo or neighboring countries. Many EU countries have adopted laws requiring companies to address in their annual reports the state of their suppliers' compliance with ILO core labor standards.

assessment of suppliers' practices, even in regimes where governments cannot be relied upon to effectively regulate companies (Montiel, Husted, and Christmann, 2012).

The results of these compliance audits provide MNCs with important information that informs their strategic decisions about supply chain partners. Adequate adherence to standards is often a prerequisite for doing business. For instance, many buyers require their suppliers to be ISO-certified to attest to their robust quality and/or environmental-management processes. Similarly, high-profile MNCs anxious to protect their brand reputations often assess the compliance of their potential and ongoing suppliers with the labor standards embodied in the MNC's code of conduct compliance as a necessary condition for commercial relationships (O'Rourke, 2003). When auditors fail to accurately assess supplier compliance, they undermine buyers' ability to make fully informed decisions about the suppliers they work with. Auditor failures can also result in catastrophic consequences, as in the case of a factory fire in Bangladesh that killed many workers even though its recent social audit results indicated that its working conditions had improved just months before the disaster (Yardley, 2012). Similarly, in 2012, just weeks after social auditors certified that a factory in Pakistan met the SA8000 working conditions standard created by a well-respected New York-based nonprofit, a factory fire killed hundreds of workers, some of whom were trapped by locked emergency exits and barred windows (Walsh and Greenhouse, 2012). In short, inaccurate assessment by social auditors can have not only business dangers, but disastrous outcomes. For these reasons, supply chain monitoring is a critical strategic question for today's MNCs.

Despite the prevalence of supply chain auditors and their central role in providing information that forms the basis of key strategic decisions, little is known about how auditors do their jobs. Much of the existing literature on private supply chain standards brackets the question

of auditor performance or implicitly assumes that the auditors who monitor compliance with these standards provide reasonably objective assessments of conditions and practices in supply chain firms. To the extent scholars address the issue of auditor performance, they have typically charged that supply chain auditors are unreliable because they are biased in favor of their paying clients (O'Rourke, 2002; Esbenshade, 2004; Heras-Saizarbitoria and Boiral, 2013). To our knowledge, however, no research rigorously has investigated what factors actually shape how supply chain auditors assess factory compliance with laws and standards.

This article seeks to fill that gap. Drawing on insights from the literatures on street-level bureaucracy in government monitoring regimes, regulatory and audit design, as well as global strategy, we theorize and investigate the factors that shape how auditors perceive and record supply chain factories' violations of private standards. We argue that auditors are not objective transmitters of actual supply chain conditions, but rather, that the information they transmit to their corporate clients is shaped by a complex web of social relationships, institutions, and identities. In particular, we explore how auditors are influenced by their ongoing relationships with the audited factories as well as the auditing team's professional experience and gender composition. By investigating these factors, we move beyond the standard focus on financial conflicts of interest to explore a variety of additional and critical factors affecting how auditors assess factory performance.

We test our hypotheses in the context of social auditing for compliance with labor standards contained in corporate supplier codes of conduct. Codes of conduct have become a ubiquitous form of private supply chain regulation. All U.S. *Fortune* 500 companies and thousands of other prominent MNCs have adopted codes governing their supply chains (McBarnet, 2007). These corporate codes are imposed through contracts with suppliers that

stipulate not only transactional terms, such as price, quantity, delivery, and quality, but also social, environmental, and human-rights standards for production processes. The substantive features of these TNC codes are highly consistent; they usually incorporate broad international consensus standards, such as the core labor standards of the International Labour Organization. The codes also typically call for the supplier to comply with domestic labor, environmental, and human-rights laws, and specifically forbid practices like child labor and prison labor even if such practices are legal in the supplier's country or prohibited only by unenforced laws (McBarnet and Kurkchiyan, 2007). We exploit a novel dataset drawn from thousands of audits for code of conduct compliance conducted in over 50 countries by one of the world's largest supply chain auditing firms.

The results of our analysis indicate the complexity of the social auditing process. We find that auditors' decisions are shaped by factors like ongoing client relationships, professional experience, gender, and gender diversity. These findings significantly broaden the prevailing understanding of the supply chain auditing process, and they suggest ways to design more effective monitoring regimes for private standards adherence in global supply chains.

LITERATURE REVIEW

A substantial body of literature addresses various issues surrounding private standards governing supply chain practices. Scholars have investigated the private standard-setting process in organizations like ISO (Wood, 2004) and the FSC (Meidinger, 2002). Many studies have documented patterns of and explanations for the adoption and diffusion of private supply chain standards including industry-specific programs and certification schemes (Hoffman, 2001; Bartley, 2007; Bartley, 2010), the ISO 9001 quality management standard (Guler, Guillen, and MacPherson, 2002; Christmann and Taylor, 2006; Terlaak and King, 2006), and the ISO 14001

environmental management standard (Christmann and Taylor, 2001; Delmas, 2002; Potoski and Prakash, 2004; King, Lenox, and Terlaak, 2005; Prakash and Potoski, 2006; Boiral, 2007; Delmas and Montiel, 2008; Delmas and Toffel, 2008). While the strategy literature has long examined corporate social performance (e.g., Waddock and Graves, 1997; Chatterji, Levine, and Toffel, 2009; Surroca, Tribo, and Waddock, 2010), scholars are increasingly examining the management of supply chains through a strategic lens (e.g., Parmigiani, 2007; Reitzig and Wagner, 2010 Alcacer and Oxley, 2013).

In addition, a significant body of research has studied outcomes for firms that adopt supply chain standards. Several studies examine how adopting various supply chain standards have affected business outcomes such as competitiveness, sales, and stock performance (Delmas, 2001; Corbett, Montes-Sancho, Kirsch, 2005; Terlaak and King, 2006; Levine and Toffel, 2010) and operational outcomes such as waste, pollution, and occupational safety (King and Lenox, 2001; Potoski and Prakash, 2005a; Levine and Toffel, 2010; Yin and Schmeidler, 2009). Others have investigated the effect of private standards on regulatory outcomes like legal compliance (Potoski and Prakash, 2005b) and working conditions (Esbenshade, 2004; Rodríguez-Garavito, 2005; Locke and Romis, 2007; Kocer and Fransen, 2009; Locke, Rising, and Pal, 2012). The strong consensus in the literature on regulatory outcomes is that private standards are most likely to positively influence performance when they are enforced through robust third-party monitoring (Potoski and Prakash, 2005b; Weil, 2005; Short and Toffel, 2010).

Although the prevailing wisdom in the supply chain standards literature assigns auditors a central role, little is known about how they perform their jobs. There has been psychological research on the cognitive motivations and biases of financial auditors (Tetlock, 1983; Glover, 1997; Hoffman and Patton, 1997; Asare, Trompeter, and Wright, 2000; Turner, 2001), but it is

not clear that these findings translate to the context of supply chain auditing. A few studies have used surveys to investigate the motivations, attitudes, and perceptions of supply chain auditors who monitor compliance with ISO 9001 quality standards (Williamson, Rogerson, and Vella, 1996; Power and Terziovski 2007). Williamson *et al*. (1996) queried auditors about their understandings of the purpose of quality control auditing, the most valuable types of evidence for establishing compliance, and the most important constituency audits were serving. Power and Terziovski (2007) surveyed auditors to assess auditing style, or how rigidly auditors adhered to formal audit protocols. Neither of these studies attempted to link auditors' attitudes, motivations or styles to their actual job performance. Most pertinent to the questions we investigate here is a qualitative study of private social auditors who monitor factories in the Dominican Republic for labor-standards compliance, which finds the auditors' work complementary to that of government labor inspectors because the two groups are subject to different incentives and political pressures (Amengual, 2010). While this study provides useful insight into factors that might shape the decisions of supply chain auditors, it also highlights the need for more comprehensive and systematic analyses like this one. Our study responds to this need and to recent calls for more empirical research into the process of supply chain auditing (Heras-Saizarbitoria and Boiral, 2013).

Although there is very little empirical research on private compliance monitors, there is a substantial body of literature on government compliance monitors, or regulatory inspectors. Econometric studies have revealed substantial heterogeneity in how government inspectors apply the rules of a given regulatory regime, but have not theorized the sources of this variation or the patterns within it (Feinstein, 1989; Feinstein, 1990; Macher, Mayo, and Nickerson, 2011; Lemley and Sampat, 2012). The literature on street-level bureaucracy investigates how the

interests and identities of individual bureaucrats, and the social institutions in which they are embedded, shape how they exercise their discretion in implementing government regulatory regimes (Lipsky, 1980/2010; Brehm and Gates, 1997; Keiser and Soss, 1998; May and Winter, 2000; Sandfort, 2000; Maynard-Moody and Musheno, 2003; Piore, 2005; Piore and Schrank, 2008; Schrank, 2009), but it has typically done so through qualitative case studies, and it has not considered whether or how its insights might extend to private monitors. Our study extends these literatures to the context of supply chain auditing for compliance with private standards for the first systematic, comprehensive investigation of the complex mix of factors that influence how these private-sector monitors apply the standards they are charged with administering.

HYPOTHESES

Although the standards embodied in supplier codes of conduct are meant to constrain the discretion of individual auditors, studies of public regulatory implementation have shown that rules are not self-executing but rather acquire meaning from the way in which they are enforced (Hawkins, 1984; Black, 1997). Street-level bureaucrats implementing government regulatory regimes must inevitably decide which rules apply to the facts they observe. Accordingly, though their work might be "rule-saturated," it is not necessarily "rule-bound" (Maynard-Moody and Musheno, 2003: 10). Like their government counterparts, private-sector supply chain auditors must decide how to apply the standards they are responsible for implementing. Below, we hypothesize several factors that influence these decisions.

Ongoing client relationships

Supply chain auditors who make repeated visits to the same factory are likely to be subject to social pressures and cognitive biases that will influence the violations they detect and

cite. Returning auditors may develop "cozy relationships" (Moore *et al*., 2006: 24) with factory management that discourage an arms-length "policing" style of enforcement in favor of one that bestows "benefit of the doubt" (Bardach and Kagan, 1982/2002). Moore *et al*. (2006), for instance, describe how financial auditors come to adopt the perspectives of their long-term clients. Our interviews with social auditors revealed similar concerns that auditors who audit the same facility multiple times risk "going native" and become an extension of the factory's management team.

Auditors' cognitive constraints also shape the number and type of violations they discover on consecutive visits to the same factory. Bounded rationality limits the number of issues that individual auditors are capable of pursuing when they conduct audits (Simon, 1947; Jones, 2001). As Chugh and Bazerman (2007: 3) have shown, individuals' "bounded awareness" causes them to "overfocus on some information and fail to use other easily available information." Specifically, individuals tend to focus on information that comports with the tacit knowledge they have gained through experience. Though tacit knowledge can be a useful resource for decision makers, "dependence on tacit knowledge can create bounds on their awareness" (Kumar and Chakrabarti, 2012: 940). Management research has found that "managers use already established knowledge to determine what they see, and they use what they already know to choose what to look for in their environment" (von Krogh and Slocum, 1994: 50). Huber and Power (1985: 172) observe that managers' "perceptual and cognitive limitations" lead to inaccurate information. Henderson and Clark (1990) demonstrate, for instance, that engineers approach new problems through the lens of their experience solving problems on previous projects, restricting their ability to identify innovative solutions. We argue that auditors will be subject to similar constraints: returning to the same factory on consecutive

visits, they are likely to focus on the set of issues they identified earlier, leading their teams to cite fewer violations. By contrast, a completely new audit team examining a factory with a fresh set of eyes may perceive a different set of issues.

> Hypothesis 1 (H1): An audit will yield fewer violations when conducted by an audit team that includes a member of the factory's previous audit team.

Auditor tenure

Scholars and activists have suggested that more experienced supply chain auditors are more effective monitors (Esbenshade, 1994; Locke, Qin, and Brause, 2007). It is not clear, however, how experience affects the number of violations cited in a given audit, because tenure on the job may exert one kind of influence on an auditor's ability to detect violations and another on the propensity to cite violations. Experience undoubtedly enhances the ability to identify violations, as would be expected and as has been documented in qualitative studies of government inspectors (Bardach and Kagan, 1982/2002). Our interviews with managers of social auditors also indicate that experience acquaints auditors with "tricks of the trade"—such as how to detect that a factory employs child labor even if child workers are not present during the audit—and that auditors exhibit "massive improvement" in their initial years on the job but that such marginal gains diminish as they learn more such techniques.

Even as more experience better equips auditors to discover violations, it is not clear that more experienced auditors will *cite* more violations. On the contrary, scholarship on government regulatory agencies has suggested that new inspectors tend to exhibit "a more policing, nit-picking attitude" than more seasoned inspectors (Bardach and Kagan, 2002: 129). Inexperienced inspectors "know too little about the industries and operations they are inspecting" and thus "lack the confidence to evaluate actual levels of risk or safety" posed by particular violations; thus they tend to go by the book and cite everything (Bardach and Kagan, 1982/2002: 129). By contrast,

experienced inspectors may decline to cite certain violations that they view as lacking the requisite level of risk and culpability (Bardach and Kagan, 1982/2002; Hawkins, 1984). Thus, we expect that auditors' violation counts will initially rise with auditor tenure, as auditors gain the experience to detect violations, but that this effect will be tempered as experienced auditors gain the confidence to exercise more discretion.

> Hypothesis 2 (H2): Audits conducted by more experienced auditor teams will yield more violations but at a decreasing rate.

Professionalization

Education and training that enhance the professionalism of supply chain auditors should promote both detection and citation of code violations. Professional training helps individuals acquire skills and expertise that should help them do their jobs better (Chen, Chang, and Lee, 2008). More professionalized auditors may also feel more obligated to cite the violations they find. Sociologists have long theorized that professionalization—specialized education and training in both the skills and the values of a particular field—is a key constraint on the exercise of individual discretion in both corporate and government bureaucracies (Scott, 1966; Larson, 1977; Abbott, 1988; Freidson, 1994). Weber (1947) argued that professionals are governed by a shared commitment to the ethics and purposes of their profession, which can motivate them to act independently of hierarchical commands; Durkheim (1893/1984) suggested that the professions help sustain community by acting as preservers and transmitters of shared values.

More recently, Lipsky (1980/2010: 201) articulated the argument that enhanced professionalism constrains the discretion of front-line workers in government bureaucracies: "[S]treet-level bureaucrats should be professionals whose relatively altruistic behavior, high standards, and self-monitoring substitute for what the society cannot dictate. Who will watch the watchmen? The watchmen will watch themselves." Scholars have also suggested that

professionalism can temper the effects of economic incentives on the behavior of employees of for-profit corporations. For instance, Parker (1999) argues that if the staff of a corporate internal compliance program were more professionalized, it might be better equipped to contest the company's profit-maximization imperatives in order to discourage wrongdoing. Although many have noted the gap between professionals' value-orientation in theory and their profit-orientation in practice (Thompson, 1967; Lipsky, 1980/2010; Gordon and Simon, 1992), professionalism remains one of the few mechanisms available to create "islands of civic virtue … in a world of generalized self-seeking" (Gordon and Simon, 1992: 235). Research has demonstrated that professionalization improved the efficacy of government labor inspectors (Piore, 2005; Schrank, 2009). Thus, we expect that teams whose auditors are more professionalized will record more violations than teams whose auditors are less professionalized.

> Hypothesis 3 (H3): Audits conducted by teams that include auditors who are more professionalized will yield more violations.

Gender

Research has suggested that, even in bureaucracies in which people are constrained by a shared set of formal rules and organizational roles, men and women "may perform their work "somewhat differently" (Eagly and Johannesen-Schmindt, 2001: 783). Several gender-based behavioral distinctions documented in the literature can influence whether social auditors discover violations and, when they do, whether they cite them. For instance, research has shown that women are more persistent than men at pursuing tasks assigned by others (Stonewater, Eveslage, and Dingerson, 1990; Spence and Buckner, 2000; Jacob, 2002), suggesting greater diligence about enforcing regulations. Research has also found that women have certain perceptual and integrative processing advantages, some of which may enhance their ability to detect violations. For example, women have been found to be more skilled than men at

interpreting the emotional content of others' expressions (Thayer and Johnsen, 2000; Killgore and Cupp, 2002; Campanella *et al*., 2004) and to be "more sensitive to subtle stimulus" (Darley and Smith, 1995: 43). Research has also found that women tend to use a more comprehensive information-processing style than men do, whereby they "attempt to assimilate all available cues" (Darley and Smith, 1995: 43). According to Gold, Hunton, and Gomaa (2009: 3):

> [W]omen tend to integrate more of the available evidential cues into their judgments, reflecting an intense level of cognitive processing. Men, on the other hand, tend to eliminate what they deem to be irrelevant cues and focus on a limited set of salient pieces of information that are relatively easy and quick to process.

Thus, women's information-gathering and processing style may equip them particularly well to perceive compliance violations in a complex factory environment and to elicit information about violations from employees.

Moreover, research suggests that women are more likely than men to cite the violations they perceive. Though all supply chain auditors are formally constrained by the rules they are charged with implementing, women in bureaucratic organizations are more likely to be strict rule-followers than men (Portillo and DeHart-Davis, 2009; Oberfield, 2010; Portillo, 2012). A long line of sociological scholarship has argued generally that "rules are a means of asserting power for the less powerful" (Portillo, 2012: 91), and that low-status members of organizations use rules as a source of authority to compensate for their lack of personal authority over others (Green and Melnick, 1950; Kanter, 1977; Thompson, 1977). We are not aware of any research on women's status in the field of supply chain auditing, but research on financial auditors and auditing firms suggests that, even as women have entered that profession in significant numbers, audit firms have maintained masculine organizational cultures that tend to devalue the contributions of female auditors (Jonnergård, Stafsudd, and Elg, 2010; Mueller, Carter, and Ross-Smith, 2011; Haynes, 2012). Our interviews with social auditors suggest that, especially in

societies where gender hierarchy is more rigid, factory managers view male auditors as more authoritative than female auditors. Empirical studies of government workers have found that women do indeed "go by the book" (Green and Melnick, 1950; Portillo and DeHart-Davis, 2009; Portillo, 2012: 90) more strictly than their male colleagues do. We believe that, taken together, this evidence suggests that supply chain auditors' gender will significantly influence whether they detect and cite violations.

> Hypothesis 4 (H4): Audits conducted by all-female teams will yield more violations than those conducted by all-male teams.

Gender diversity

Supply chain auditing teams are not necessarily all-male or all-female. In the organizational literature on teams, there is significant debate about the effects of diversity, including gender diversity, on team performance (Joshi and Roh, 2009; Phillips *et al*., 2012). We expect that, in the context of supply chain auditing, gender diversity will enhance team performance because of complementary perceptual styles and interpersonal dynamics.

First, women's and men's different perceptual styles may cause them to identify different types of violations, enabling mixed-gender teams to find more violations in the aggregate. This kind of perceptual complementarity should be particularly valuable in the supply chain context, where audit teams are charged with eliciting information from a diverse set of employees and managers. Research has demonstrated that, for a variety of reasons, "diversity in groups increases the likelihood that there will be access to different information in a group" (Phillips *et al*., 2012: 161). Our interviews with social auditors indicate that factories subjected to social audits tend to have predominantly female workers and male managers, and that the female workers are more likely to communicate openly with female auditors than with their male counterparts. All-female audit teams, however, might have more difficulty than teams including

male auditors getting access to factory managers, who are nearly always male and, in the words of one interviewee, "for cultural reasons, may find it difficult…to open up to women."

Second, research has identified mechanisms whereby gender diversity can create interpersonal dynamics within teams that raise overall performance. For instance, studies have shown that people on socially diverse teams tend to prepare more thoroughly and to think through a broader range of issues than do those on non-diverse teams (Loyd *et al*., 2013). Fenwick and Neal report the superior performance of gender-diverse teams at management-simulation exercises, which they attribute to the "mix of male and female operating, decision-making and leadership styles" (2001: 217). Furthermore, men on mixed teams may try harder if they perceive that they are performing less well than the women. Studies have shown generally that lower-performing team members perceive their higher-performing counterparts as standards against which to judge their own efforts, and tend to raise their "usual performance levels in order to match or beat the stronger performers" (Weber, Wittchen, and Hertel, 2009: 732; Collins, 2000; Lount, Jr., and Phillips, 2007). Weber, Wittchen, and Hertel (2009) demonstrate that men are particularly prone to such competitive behavior. Male auditors may therefore feel compelled to match or exceed the violation citation rates of their female teammates.

Although some conflicting evidence suggests that gender diversity can undermine team performance in certain circumstances, a recent meta-study found that gender diversity is particularly likely to enhance performance in service industries, where team members interact directly with clients (Joshi and Roh, 2009). Because supply chain auditing is a service industry, and auditors engage in substantial interpersonal interaction with clients, we expect that gender diversity will positively influence team performance.

> Hypothesis 5 (H5): Audits conducted by gender-diverse audit teams will yield more violations than those conducted by single-gender teams.

DATA AND MEASURES

Empirical context and sample

To test our hypotheses, we obtained data associated with thousands of code-of-conduct audits conducted by one of the world's largest social-auditing companies between 2004 and 2009. During that period the company, which required anonymity as a condition of sharing its data with us, conducted social audits in over 60 countries and had more than a decade's experience. The company employed several hundred people in many countries; they spoke over 30 languages.

The dataset contains audit results and information about each audited factory, including its country and a unique identifier, but anonymity was preserved by not revealing to us the factories' names. Similarly, the dataset included characteristics of and unique identifiers for the auditors who conducted each audit but not their names. The data also included the country and a unique identifier for the client on whose behalf each audit was conducted but not the clients' names. Our estimations are based on the 16,795 audits of 5,819 factories (in 66 countries) for which data were available about all the measures described below and among factories that were audited at least twice during the sample period (a technical requirement owing to our models being estimated with factory-level fixed effects, described below). The industry composition of our sample is reported in Table 1. The most common industries of the audited factories in our dataset are garments, accessories, electronics, and toys.

[Insert Table 1 about here]

In nearly all cases, brands determined which factories would be audited, and whether it would pay for the audit or whether it would require the factory or its agents to pay. Our interviews indicated that this decision was not driven by the factory's managerial attitude,

violation rate, or improvement rate. (Factories sometimes sought and paid for audits when they sought to become certified to a third-party standard such as SA8000. As described later, our results are robust to omitting from the estimation sample the very small proportion of audits that used third-party protocols.)

Dependent variable

We measure the extent to which factories adhere to codes of conduct as the *number of violations* in each audit, obtained from the social auditing firm's database. We include only the types of code-of-conduct violations that the social auditing firm confirmed apply to factories in all industries and that are interpreted by auditors in the same way in all countries, namely violations of rules for child labor, forced or compulsory labor, working hours, occupational health and safety, minimum wage, treatment of foreign workers and subcontractors, and disciplinary practices.[2] In the course of each audit, the auditors code a common set of dichotomous indicators (violation or no violation) in each of these seven categories. For example, the occupational health and safety category consists of seven indicators pertinent to emergency preparedness (blocked or locked aisles or exits, inadequate first-aid supplies, insufficient emergency exits, lack of emergency lighting, lack of employee emergency training, lack of an evacuation plan, and unmarked aisles), five indicators of fire safety, eight related to toilets, and eight related to the work floor.

Independent variables

To identify the potential for auditors' career concerns to influence their behavior, we identified instances in which a member of a factory's focal audit team had been a member of a

[2] We exclude other categories that, according to our auditor interviews, applied only to factories in particular industries or that were interpreted differently in different countries: the right of association, the right to organize and bargain collectively, legal client regulation, dormitory conditions, and canteen violations.

prior audit team at the same factory. Specifically, we coded *previous auditor* as 1 when at least one member of the focal audit team had participated in one of the factory's previous audits during the sample period and 0 otherwise.

We measure the experience of the auditors on each team as their years of service at the social-auditing company. Using data from the social auditing firm's database, we calculated *maximum tenure* as the highest number of years that any member of the audit team had worked at the company. (Using average tenure, rather than maximum tenure, as an alternative measure of the audit team's experience yielded nearly identical results.)

We measure the professionalism of the audit team in two ways. Because one important source of professionalization is "standardized formal training in universities" (Lipsky, 1980/2010: 201), we created *graduate education* as a dichotomous variable coded 1 when at least one member of the audit team possessed a graduate degree and 0 otherwise. We focused on graduate education because nearly all auditors in our dataset possessed a bachelor's degree.[3] We also created *auditing skills training* as the highest number of the social auditor's training courses that any audit team member had completed. These training courses teach audit skills, such as how to interpret national labor laws and how to detect payroll manipulation that might indicate wage violations. (Using the average rather than the maximum number of training courses as an alternative measure of the audit team's training yielded nearly identical results.)

Because the social auditing firm's database indicated that 97 percent of the audit teams in our sample were all-female, all-male, or evenly divided, we measure gender composition by creating three dummy variables: *all-female audit team, all-male audit team,* and *mixed-gender audit team*.

[3] We coded graduate education as a dichotomous variable rather than a continuous measure to better reflect the near-binary distribution in our sample: 87 percent of the audit teams had no members with a master's degree, 7 percent had all members with a master's degree, and a mere 6 percent had an intermediate configuration.

Control variables

Prior research indicates that training can influence the stringency of government regulators (Macher, Mayo, and Nickerson, 2011), and that it might also influence an audit team's ability to detect and report violations. We thus control for two types of training. Using the social auditing firm's database, we calculated the proportion of each audit team that had undergone *certification training*—that is, training on the standards and protocols of a particular certification regime, such as SA8000—and the proportion that had undergone *brand training*, or training provided by the brand on its corporate responsibility program and procedures.

We control for auditors' age to ensure that the effects of auditor tenure can be attributed to experience on the job rather than the life-cycle effects posited by human capital theory (Diamond, Jr., 1984), which predicts "an inverse U-shaped relationship between productivity and age" (Teitelbaum, 2006: 166). Because the social auditing firm provided a five-year age-range category (for example, 20–24 years old) for each auditor to keep precise ages confidential, we created a proxy for the audit team's average age. Specifically, we calculated the midpoint for each category and then created *average age* as the average of the oldest and youngest age-range categories represented on a given audit team. (Using the age of the oldest audit team member rather than the average yielded nearly identical results.)

We control for whether the audit used a third-party protocol because such audits might limit an auditor's discretion. We created a dummy variable to indicate whether an audit used a *third-party protocol* such as that of the Business Social Compliance Initiative (BSCI), the Initiative Clause Sociale (ICS), the Sedex Members Ethical Trade Audit (SMETA), the International Council of Toy Industries (ICTI), or the Worldwide Responsible Accredited Production (WRAP).

We also control for whether an audit is unannounced or pre-announced. The latter practice provides several weeks of advance notice, giving the factory time to try to remedy problems before the auditor arrives, which could result in fewer violations to discover and report. In contrast, a factory does not know the date of an unannounced audit, though it may have been given a window of several weeks. We created a dummy variable *unannounced audit* coded 1 for an unannounced audit and 0 for a pre-announced audit.

Because research has indicated that financial conflicts of interest created by client fees undermines auditors' and inspectors' ability to police corporate misconduct (Cantor and Packer, 1994; Bazerman, Morgan, and Loewenstein, 1997; Oh, 2004; Moore *et al.*, 2006; Partnoy, 2006; Dallas, 2011; Estlund, 2012; Pierce and Toffel, 2013; Manns, 2013), we control for which entity paid for each social audit. Based on data from the social auditing firm's database, we created two dichotomous variables. *Audit paid for by factory or agent* identifies audits with the potential for financial conflict of interest as described earlier. This variable is coded 1 for audits paid for by the audited factory, agents, vendors, or licensees, and 0 for audits paid for by the brand.[4] *Audit paid for by brand* is coded in the opposite manner.

We also include a series of dummy variables to control for the *number of auditors* who conducted each audit (two through five, with one as the omitted category), which is a direct function of the size and complexity of the audited factory.

We created a dummy variable *re-audit* to distinguish routine audits from re-audits, because re-audits tend to focus on domains where violations were identified in a previous audit

[4] We combined these categories of payers because prior research and our auditor interviews suggest that, in our empirical context, the financial incentives of factories and these intermediaries are closely aligned. In developing economies, intermediaries' role is to promote exports by domestic manufacturers by identifying new markets for their goods and services (Ellis, 2010) and by reducing transaction-cost barriers to export (Ahn, Khandelwal, and Shang-Jin, 2011). Our results are robust to an alternative specification in which we include two dummies that control for audits paid by factories as distinct from audits paid by agents, vendors, or licensees.

and thus are apt to yield fewer overall violations than routine audits. We also include a series of dummies to indicate a factory's audit sequence—its *second audit, third audit,* and so on, through *sixth-or-higher audit* (because only 5 percent of the audits in our sample were a factory's seventh or higher audit), with a factory's first audit as the omitted category—to control for the possibility that factories address issues as they are exposed, which could result in subsequent audits yielding fewer violations. (Using an *audit sequence* counter variable and its square rather than this set of dummies yielded nearly identical results.)

To capture domestic institutional factors that could influence a factory's compliance with codes of conduct (Toffel, Short, and Ouellet, 2013), we control for several country-level governmental, economic, and civil-society attributes. We measure the average economic development level of an audited factory's country as its annual *per-capita gross domestic product (GDP)* in 2005 dollars, calculated by the U.S. Department of Agriculture's Economic Research Service (obtained from http://www.ers.usda.gov). To reduce skew, we used the log of *per-capita GDP* in our models. We measure the extent to which the government of the audited factory's country fosters a regulatory environment that promotes economic development via the annual *regulatory quality* metric calculated by the World Bank's Worldwide Governance Indicators project (obtained from http://data.worldbank.org/data-catalog/worldwide-governance-indicators). The World Bank created this index to capture "perceptions of the ability of the government to formulate and implement sound policies and regulations that permit and promote private sector development" (World Bank, 2013).

We measure the extent of *press freedom* in the audited factory's country via the annual Press Freedom Index produced by Reporters without Borders (obtained from http://en.rsf.org). This index incorporates the extent to which journalists face direct and indirect threats—including

imprisonment, physical attacks, censorship, and self-censorship—and the number of journalists detained, murdered, physically attacked, or threatened. We created *press freedom* by reverse-coding the Press Freedom Index, so that a higher score represents more press freedom, and then rescaling the result to range from 0 to 1.

Summary statistics and correlations are reported in Tables 2 and 3.

[Insert Tables 2 and 3 about here]

EMPIRICAL MODEL AND RESULTS

Our empirical model includes all independent and control variables described above, as well as three sets of fixed effects. We include fixed effects for each audited factory to control for all time-invariant factory characteristics that might affect its violation rate, such as its size, industry, factory age, and its national institutional context.[5] A series of dummies for the year in which the audit was conducted controls for overall temporal trends. We also include fixed effects for the headquarters country of the brand on whose behalf each audit was conducted. This controls for the possibility that consumers and activist groups in brand countries vary in their level of concern for and attentiveness to supply chain conditions, which might in turn affect how much pressure buyers exert on their auditors to audit stringently. These fixed effects also control for all other time-invariant differences between brand headquarters countries' national institutional contexts.

Our identification strategy is based on the fact that the social auditor process of assigning auditors to audit teams is unrelated to our independent variables, and thus it does not pose a source of endogeneity that should bias their coefficients. Specifically, our interviews with the social-auditing firm indicated that auditors were assigned to audit teams based on three factors:

[5] Because factory-level fixed effects in our model absorb the time-invariant portion of factory country-level variables, *per-capita gross domestic product (GDP), regulatory quality, and press freedom* effectively control for within-country temporal variation in their effect on factory violation rates.

(1) the audit team needed the pertinent language skills to communicate with management and workers; (2) auditor availability given their other scheduled audit commitments; and (3) at least one of the auditors had to have sufficient auditing skills training to qualify as a senior auditor.

We estimate the model using Poisson regression with robust standard errors, and report our results in Column 1 of Table 4. Negative-binomial regression with conditional fixed effects yields nearly identical results, which indicates that our results are not sensitive to estimation technique.

[Insert Table 4 about here]

Results

The significant negative coefficient on *previous auditor* ($\beta = -0.04$; $p = 0.03$; IRR = 0.96) indicates that audits yield 4 percent fewer violations when an audit-team member had participated in a prior audit at the same factory, a finding that supports Hypothesis 1. The average marginal effect of -0.3 indicates that, compared to the sample average of 6.5 violations, an audit by a team that included an auditor who had participated in a previous audit at the same factory would yield 6.2 violations.

The audit team's maximum tenure has a significant positive coefficient ($\beta = 0.07$; $p < 0.01$) and its square term has a significant negative coefficient ($\beta = -0.004$; $p < 0.01$), implying that the number of violations increases as tenure increases but at a diminishing rate, which supports Hypothesis 2. This relationship is illustrated in Figure 1, which graphs average predicted values at varying levels of the audit team's maximum experience.

[Insert Figure 1 about here]

Our results are mixed with respect to professionalized auditors. The coefficient on *graduate education* is positive as predicted but not significant, and thus yields no evidence that

audits conducted by audit teams with more formal education yielded significantly more violations. Audits did yield significantly more violations when conducted by more professionalized auditors as measured by *auditor training* ($\beta = 0.02$; $p < 0.01$; IRR = 1.02). This result is not driven by auditor age or tenure because we control for these factors. The average marginal effect indicates that each additional training course (pursued by the most highly trained member of the audit team) is associated with an additional 0.14 violations. In other words, an audit team whose most highly trained member had taken nine training courses would, on average, cite one more violation than a team whose most highly trained member had taken two training courses. Jointly, these results yield some support for Hypothesis 3's prediction that audits conducted by more professionalized auditors yield more violations, but only when professionalization is measured by specific training rather than broader education.

The gender composition of the audit team is also significantly associated with the number of violations reported. Audits conducted by all-female auditor teams yield 6 percent more violations than those conducted by all-male teams (the baseline) ($\beta = 0.05$; $p < 0.01$; IRR = 1.05), a finding that supports Hypothesis 4. The average marginal effect indicates that audits conducted by all-female teams yield 0.35 more violations than those conducted by all-male teams (the baseline category).

Audit teams composed of both men and women yield on average 7 percent more violations—or nearly half a violation more—than all-male teams (the baseline) ($\beta = 0.07$; $p < 0.01$; IRR = 1.07). Mixed-gender teams also yield slightly more violations than all-female audit teams (Δ AME = 0.1), but the difference is not statistically distinguishable ($\chi^2 = 0.55$; $p = 0.46$). These results provide partial support for Hypothesis 5.

We also find that subsequent audits of a given factory yield significantly fewer violations.

The coefficients on *second inspection* through *sixth or more inspection* are all negative and statistically significant, indicating that, on average, factories' successive audits yield fewer violations than their first one. Wald tests comparing these coefficients indicate that, on average, each successive audit at a factory yields significantly fewer violations than its predecessor. Average marginal effects indicate that, on average, a factory's second audit yields nearly one fewer violation than its initial audit during our sample period ($\beta = -0.15$; $p < 0.01$; AME = -0.9), its third audit yields 1.3 fewer violations than its second audit (AME = -2.2, a statistically significant decline: Wald $\chi^2 = 114$; $p < 0.01$), and its fourth audit yields 0.6 fewer violations than its third audit (AME = -2.8, a statistically significant decline: Wald $\chi^2 = 20$; $p < 0.01$). This relationship is also apparent in the summary statistics depicted in Figure 2.

[Insert Figure 2 about here]

Consistent with the practice of assigning more auditors to larger factories, which are likely to generate more violations, we find that audits conducted by more auditors yield significantly more violations. We find no evidence that the number of violations cited in a factory's audit varied with the audit team's certification training, brand training, or average age, or whether the audit was conducted according to a third-party protocol. Our point estimate indicates that unannounced audits yielded slightly more violations than announced audits at a given factory (AME = 0.2), but the difference was outside conventional significance levels ($p = 0.15$).

Audits paid for by factories or agents yielded 8 percent fewer violations than audits paid for by brands, the baseline category ($\beta = -0.08$; $p < 0.01$; IRR = 0. 92). The average marginal effect (AME) indicates that, on average, audits yield 0.6 fewer violations when the factory or agent pays than when the brand pays, or a decline from the sample average of 6.5 violations to

5.9 violations.

Audits yielded fewer violations in countries with greater economic activity, as measured by *per capita GDP* (β = -0.62; p = 0.02; AME = -4.0), and those with greater *press freedom* (β = -0.51; p = 0.02; AME = 3.3). Our point estimate indicates fewer violations at factories in countries with higher *regulatory quality*, but the relationship was not statistically significant (β = -0.18; p = 0.22).

Robustness tests

As noted earlier, the results were nearly identical to our primary results when we estimated the model using negative binomial regression instead of Poisson regression and when we used alternative measures of the audit team's experience (mean instead of maximum tenure), training (average rather than maximum number of training courses), and age (the age of the oldest audit team member rather than the average of the oldest and youngest members' ages), and an alternative approach to controlling for the factory's audit sequence (a counter and its square instead of dummies).

We also estimated our primary model on various subsamples to assess the extent to which our results were driven by certain types of audits. Column 2 of Table 4 reports estimates after excluding the 210 audits performed for clients whose audit teams were always all-female, in case that pattern reflected a client policy that might bias our primary results. Column 3 reports estimates based on the 10,648 audits conducted by audit teams of at least two members to ensure that our results were not being driven by audits conducted by a single auditor. Column 4 reports estimates of the model after excluding the 751 audits conducted according to third-party protocols, in case the influence of such protocols on the discovery or reporting of violations is not adequately controlled for with the dummy variable used in our main specification, and

because factories themselves might have chosen to pursue those particular audits. Column 5 reports results for the subsample of 9,266 audits that exclude each factory's first inspection in our sample, because some of those initial inspections might have been pre-assessments of factories that clients had not yet engaged and our hypothesized relationships might operate differently in such circumstances. Overall, our results are quite robust across these subsamples. The sign and magnitude of all hypothesized variable coefficients are very similar to our main results.

DISCUSSION

Our research has theorized and tested several social factors that shape supply chain auditors' ability to identify and report working conditions at factories that violate suppler codes of conduct. Beyond the financial conflicts of interest that have to date been the focus of research on auditor decision making, we find that their decisions are shaped by social factors, including their experience, professional training, and gender; the gender diversity of their teams; and their repeated interactions with those whose actions they are paid to assess. These findings contribute to several literatures and suggest strategies for designing private monitoring regimes to provide companies with more reliable strategic information about their supply chain partners.

Contributions to auditing and gatekeeping research

Although much is known about the adoption, diffusion, and outcomes of supply chain standards and codes of conduct that require auditing, auditors' practices have largely remained a black box. Our research addresses this by illuminating how auditing practices implemented on the ground, at the micro level, are influenced by several key auditor characteristics.

Prior literature exploring auditor bias has focused on economic incentives and conflicts of interest that arise when those being audited are paying for the audits. Our analysis indicates

that while economic incentives play a role, private-sector auditor behavior is significantly influenced by social institutions, identities, and relationships. Our finer-grained picture of the factors that shape how auditors do their jobs suggests strategies for moderating potential bias in reporting, thus enhancing the reliability of information generated by private supply chain monitors. Audit designers seeking to build robust regulatory regimes should consider the auditors' characteristics and relationships that we found to significantly influence auditors' decisions.

Our findings about the factors that shape supply chain auditing practices should likewise inform the broader literature on private gatekeepers like accountants and credit rating agencies that monitor corporate compliance with financial regulations. There has been significant academic and policy interest in these private gatekeepers since their failures to detect and reveal corporate wrongdoing led to corporate scandals and financial meltdowns in the early twenty-first century (Bratton, 2002; Partnoy, 2004). However, similar to the auditing literature, research on gatekeepers has focused almost exclusively on how they are affected by economic conflicts of interest (Manns, 2013; Cantor and Packer, 1994; Bazerman, Morgan, and Loewenstein, 1997; Oh, 2004; Moore *et al*., 2006; Partnoy, 2006; Dallas, 2011; Goldberg, 1988; Schwarcz, 2002; Hill, 2004). Our study suggests the need to look beyond economic incentives in structuring effective gatekeeping regimes.

Contributions to research on public-sector monitors

By drawing on research on street-level bureaucracy in government regulatory agencies to predict the behavior of private-sector supply-chain auditors, our study initiates a productive dialogue between literatures on public- and private-sector monitoring that until now have each missed important insights offered by the other. Calls for insight into the micro-level processes of

private supply-chain auditing (Heras-Saizarbitoria and Boiral, 2013) have overlooked the extensive literature on street-level policy implementation by government monitors, and street-level bureaucracy research has largely ignored front-line implementation by private-sector monitors who play an increasingly important role in regulating corporate conduct. Our study extends both literatures by elaborating micro-level implementation processes in the context of private-sector auditing.

In addition, our findings extend existing economic research on government regulatory inspectors investigating how experience affects stringency. Several studies have documented less stringency among more-experienced government monitors including patent examiners (Lemley and Sampat 2012) and U.S. Food and Drug Administration inspectors of pharmaceutical manufacturing plants (Macher, Mayo, and Nickerson, 2011). But these studies tell only one side of the story, as they fail to document the initial gains to experience that our findings reveal. Our findings similarly contribute to economic literature on human capital theory (Diamond, Jr., 1984), which predicts "an inverse U-shaped relationship between productivity and age" (Teitelbaum, 2006: 166), by demonstrating that these effects may be driven by experience rather than age.

Contributions to research on transnational business regulation

Our research also advances the literature on transnational business regulation. Supply chain auditing has become an important component of international regulatory strategies that seek to address the social and environmental risks of business activities beyond the reach of state governments (Braithwaite and Drahos, 2000; Scott, 2012). Private labeling regimes like the Forest Stewardship Council, the Marine Stewardship Council, and Fair Trade rely on private, third-party auditors to certify compliance with their respective standards. International,

intergovernmental institutions like the United Nations have encouraged supply chain auditing by requesting that TNCs conduct "due diligence" to ensure their suppliers' compliance with international human rights norms (Ruggie, 2008; Shamir, 2005; Kamatali, 2012). Many national domestic regulators have followed suit, requiring that MNCs conduct due diligence and disclose practices in their supply chains. The efficacy and legitimacy of these efforts depends in large part on the credibility of supply chain monitoring, and our study responds to calls for more empirical research on the key actors implementing private transnational regulation (Büthe, 2010). While our findings of auditor heterogeneity support those who question auditor independence and objectivity (Power, 1997; Boiral and Gendron, 2011), our identification of several systematic determinants of auditor heterogeneity suggests how companies and policymakers can improve the validity of audits.

Implications for managers

Our findings yield several practical insights that can inform the way audits are structured. Our finding that auditors tend to cite fewer violations at factories where they have ongoing relationships provides empirical support to those who have advocated auditor rotation to prevent them from becoming beholden to or captured by long-term clients (Moore *et al*., 2006; U.S. Public Company Accounting Oversight Board, 2011). That auditors returning to the same factory might consistently deemphasize some areas (and thus inadvertently overlook violations) should also encourage auditing firms, and the brands that hire them, to ensure that auditors are periodically rotated.[6] Our findings also highlight the importance of training auditors. While auditors with higher educational credentials did not find significantly more violations than less educated peers, those with more audit-specific training did. This suggests the importance of

[6] A few auditing schemes have also explicitly stipulated term limits for auditing companies, such as California's Greenhouse Gas regulation requiring regulated entities to change verification companies every six years .

developing robust training systems for private-sector monitors. Finally, our findings suggest that auditing teams can benefit from gender diversity and experience, although managers should be aware that the benefits of the latter attenuate over time.

Limitations and future research

We acknowledge several limitations of our study and suggest a number of promising avenues for future research. Given the nature of our large quantitative study, we are unable to identify the precise mechanisms by which the factors we identify influence individual auditor decisions. We encourage future research to investigate the social processes underlying these outcomes.

Our many discussions with social auditors, including those employed by the firm that provided our data and employees of competing firms, yielded no reason to suspect that endogeneity concerns are driving our results. For example, these discussions indicated that assignments were driven largely by the goal of ensuring that each audit had a lead auditor (qualified by training hours) and that all auditors were assigned based on language skills, industry expertise, and availability. These discussions also indicated that the brand determines which factories are audited; this practice obviates the risk of a potential selection effect whereby better-than-average or worse-than-average factories might opt in to being audited or to paying for their own audits, as occurs in some voluntary environmental programs (King and Toffel, 2009). Despite these assurances, we cannot entirely rule out the possibility that omitted variables are correlated with our independent variables and violation rates, and we encourage future research to engage in randomized field experiments (e.g., Hainmueller, Hiscox, and Sequeira, 2011).

Also, our findings relating to gender and gender diversity may be influenced by the

gender composition of the workforce at audited factories. Although we do not have data on the employee demographics of the audited factories, our interviews with social auditors and available meta-data suggest that women dominate the workforce in export-intensive industries such as garments, textiles, and electronics, which account for most of our sample (Jenkins, Esquivel, and Larrían, 2001; Dejardin and Owens, 2009; Kuncoro, 2011). Future research could explore how auditors' decisions are influenced by the interaction of the gender composition of the audit team and that of the organization it monitors.

Future research can also explore how auditors' decisions are influenced by various short- and long-term organizational structures and incentives. For instance, differing compensation systems may influence the extent to which supply chain auditors' decisions are shaped by economic incentives and other factors. In addition, field experiments might shed light on which types of technical and managerial training most improve the objectivity of auditors' decisions. More broadly, it is important for future research to investigate whether our findings are generalizable to different types of private gatekeepers, like financial auditors, credit ratings agencies, and attorneys: do these professionals respond similarly to economic incentives, professional obligations, and social pressures in their environments? Finally, future studies that directly compare the implementation practices of private-sector monitors like social auditors and public monitors like government inspectors would promote better understanding of the efficacy and legitimacy of both.

CONCLUSION

Despite the growing importance of private supply chain auditors to strategic corporate decisions and in public and private regulatory regimes, they have seldom attracted academic attention. Our investigation of supply chain auditing practices at thousands of factories around

the world reveals several social factors that influence auditors' decisions. More broadly, our work contributes to literatures on private supply chain monitoring, private gatekeeping, street-level bureaucracy and transnational business regulation and highlights opportunities to improve the design and implementation of auditing regimes that rely on private-sector monitors and to enhance their accountability and legitimacy.

REFERENCES

Abbott, A. 1988. *The System of Professions: An Essay on the Division of Expert Labor*. Chicago: University of Chicago Press.

AFL-CIO. 2013. *Responsibility Outsourced: Social Audits, Workplace Certification and Twenty Years of Failure to Protect Worker Rights.* Washington, D.C.: American Federation of Labor-Congress of Industrial Organizations (AFL-CIO).

Ahn, J., A. K. Khandelwal, and S. Wei. 2011. The role of intermediaries in facilitating trade. *Journal of International Economics* 84(4): 73–85.

Alcacer, J. and J. Oxley. 2013. Learning by supplying. *Strategic Management Journal*. 35(2): 204–223.

Amengual, M. 2010. Complementary labor regulation: The uncoordinated combination of state and private regulators in the Dominican Republic. *World Development* 38(3): 405–414.

Asare, S. K., G. M. Trompeter, and A. M. Wright. 2000. The effect of accountability and time budgets on auditors' testing strategies. *Contemporary Accounting Research* 17(4): 539–560.

Bansal, P. 2005. Evolving sustainably: A longitudinal study on corporate sustainable development. *Strategic Management Journal* 26: 197–218.

Bardach, E., and R. A. Kagan. 1982/2002. *Going by the Book: The Problem of Regulatory Unreasonableness*. New Brunswick, NJ: Transaction.

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. 2010. Transnational private regulation in practice: The limits of forest and labor standards certification in Indonesia. *Business and Politics* 12(3): Article 7.

Bazerman, M. H., K. Morgan, and G. F. Loewenstein. 1997. The impossibility of auditor independence. *MIT Sloan Management Review* 38(4): 89–94.

Berrone, P, A. Fosfuri, L. Gelabert, and L. Gomez-Maria. 2013. Necessity as the mother of 'green' inventions: Institutional pressures and environmental innovations. *Strategic Management Journal* 34: 891–909.

Black, J. 1997. *Rules and Regulators*. Oxford, UK: Clarendon Press.

Boiral, O. 2007. Corporate greening through ISO 14001: A rational myth? *Organization Science* 18: 127–146.

Boiral, O., and Y. Gendron. 2011. Sustainable development and certification practices: Lessons learned and prospects. *Business Strategy and the Environment* 20: 331–347.

Braithwaite, J., and P. Drahos. 2000. *Global Business Regulation*. Cambridge, UK: Cambridge University Press.

Bratton, W. W. 2002. Enron and the dark side of shareholder value. *Tulane Law Review* 76: 1275–1361.
Brehm, J., and S. Gates. 1997. *Working, Shirking, and Sabotage*. *Bureaucratic Response to a Democratic Public*. Ann Arbor, MI: University of Michigan Press.
Büthe, T. 2010. Global private politics: A research agenda. *Business and Politics* 12(3): Article 12.
Campanella, S., M. Rossignol, S. Mejias, F. Joassin, P. Maurage, D. Debatisse, R. Bruyer, M. Commelinck, and J.-M Guerit. 2004. Human gender differences in an emotional visual oddball task: An event-related potentials study. *Neuroscience Letters* 367(1): 14–18.
Cantor, R., and F. Packer. 1994. The credit rating industry. *Federal Reserve Bank of New York Quarterly* 19: 1–26.
Chatterji, A. K., D. I. Levine, and M. W. Toffel. 2009. How well do social ratings actually measure corporate social responsibility? *Journal of Economics & Management Strategy* 18(1): 125–169.
Chen, Y., B. Chang, and C. Lee. 2008. The Association between continuing professional education and financial performance of public accounting firms. *International Journal of Human Resource Management* 19(9): 1720–1737.
Christmann, P., and G. Taylor. 2001. Globalization of the environment: determinants of firm self-regulation in China. *Journal of International Business Studies* 32: 439–458.
Christmann, P., and G. Taylor. 2006. Firm self-regulation through international certifiable standards: Determinants of symbolic versus substantive implementation. *Journal of International Business Studies* 37: 863–883.
Chugh, D., and M. H. Bazerman. 2007. Bounded awareness: what you fail to see can hurt you. *Mind and Society* 6: 1–18.
Collins, R. L. 2000. Among the better ones: Upward assimilation in social comparison. In J. Suls and L. Wheeler (eds.), *Handbook of Social Comparison: Theory and Research:* 159–172. New York: Plenum.
Corbett, C. J., M. J. Montes-Sancho, and D. A. Kirsch. 2005. The financial impact of ISO 9000 certification in the United States: An empirical analysis. *Management Science* 51(7) 1046–1059.
Dallas, L. L. 2011. Short-termism, the financial crisis, and corporate governance. *Journal of Corporation Law* 37(2): 264–363.
Darley, W. K., and R. E. Smith 1995. Gender differences in information processing strategies: An empirical test of the selectivity model in advertising response. *Journal of Advertising* 24(1): 41–56.
Dejardin, A. K., and J. Owens. 2009. *Asia in the Global Economic Crisis: Impacts and Responses from a Gender Perspective*. Bangkok: International Labour Organization Regional Office for Asia and the Pacific.
Delmas, M. 2001. Stakeholders and competitive advantage: The case of ISO 14001. *Production and Operations Management* 10: 343–358.
Delmas, M. 2002. The diffusion of environmental management standards in Europe and in the United States: An institutional perspective. *Policy Sciences* 35: 91–119.
Delmas, M. A., and I. Montiel. 2008. The diffusion of voluntary international management standards: Responsible care, ISO 9000 and ISO 14001 in the chemical industry. *Policy Studies Journal* 1: 65–93.
Delmas, M. and M.W. Toffel. 2008. Organizational responses to environmental demands:

Opening the black box. *Strategic Management Journal* 29(10): 1027–1055.
Diamond, A. M., Jr. 1984. An economic model of the life-cycle research productivity of scientists. *Scientometrics* 6: 189–196.
Duhigg, C., and D. Barboza. 2012. The iEconomy, Part 2: A punishing system. In China, the human costs that are built into an iPad. *New York Times*, January 26: A1.
Durkheim, E. 1893/1984. *The division of labor in society*, trans. W. D. Halls. New York: Free Press.
Eagly, A. H., and M. C. Johannesen-Schmindt. 2001. Leadership styles of women and men. *Journal of Social Issues* 57(4): 781–797.
Ellis, P. D. 2010. International trade intermediaries. In *Wiley International Encyclopedia of Marketing*, DOI: 10.1002/9781444316568.wiem06040.
Esbenshade, J. 1994. *Monitoring sweatshops: Workers, Consumers, and the Global Apparel Industry*. Philadelphia: Temple University Press.
Esbenshade, J. 2004. *Monitoring Sweatshops: Workers, Consumers and the Global Apparel Industry*. Philadelphia, PA: Temple University Press.
Estlund, C. 2012. Enforcement of private transnational labor regulation: a new frontier in the anti-sweatshop movement? In F. Cafaggi (ed.), *Enforcement of transnational regulation: Ensuring compliance in a global world:* 237–262. Cheltenham, UK: Edward Elgar.
Feinstein, J. S. 1989. The safety regulation of U.S. nuclear power plants: Violations, inspections, and abnormal occurrences. *Journal of Political Economy* 97: 115–154.
Feinstein, J. S. 1990. Detection controlled estimation. *Journal of Law and Economics* 33(1): 233–276.
Fenwick, G. D., and Neal, D. J. 2001. Effect of gender composition on group performance. *Gender, Work and Organization* 8(2): 205–224.
Freidson, E. 1994. *Professionalism Reborn: Theory, Prophecy, and Policy*. Chicago: University of Chicago Press.
Galland, A. 2010. *Toward a Safe, JustWorkplace: Apparel Supply Chain Compliance Programs*. San Francisco: As You Sow.
Glover, S. M. 1997. The influence of time pressure and accountability on auditors' processing of nondiagnostic information. *Journal of Accounting Research* 35(2): 213–226.
Gold, A., J. E. Hunton, and M. I. Gomaa. 2009. The impact of client and auditor gender on auditors' judgments. *Accounting Horizons* 23(1): 1–18.
Goldberg, V. P. 1988. Accountable accountants: Is third party liability necessary? *Journal of Legal Studies* 17(2): 295–312.
Gordon, R. W., and W. H. Simon. 1992. The redemption of professionalism? In R. L. Nelson, D. M. Trubek, and R. L. Solomon (eds.), *Lawyers' Ideals, Lawyers' Practices: Transformations in the American Legal Profession:* 230–257. Ithaca, NY: Cornell University Press.
Green, A. W., and E. Melnick.1950. What has happened to the feminist movement? In A. W. Gouldner (ed.), *Studies in Leadership:* 277–302. New York: Harper and Brothers.
Greenhouse, S. 2013. $40 Million in Aid Set for Bangladesh Garment Workers. *New York Times*: December 24: B1.
Guler, I., M. F., and J. M. MacPherson. 2002. Global competition, institutions, and the diffusion of organizational practices: The international spread of ISO 9000 quality certificates. *Administrative Science Quarterly* 47: 207–232.
Hainmueller, J., M. J. Hiscox, and S. Sequeira. 2011. Consumer demand for the Fair Trade label: Evidence from a field experiment. Research Paper No. 2011-9B, MIT Political Science

Department.
Hawkins, K. 1984. *Environment and Enforcement: Regulation and the Social Definition of Pollution*. London: Oxford University Press.
Haynes, K. 2012. Body beautiful? Gender, identity and the body in professional services firms. *Gender, Work & Organization* 19(5): 489–507.
Henderson, R. M., and K. B. Clark. 1990. Architectural innovation: the reconfiguration of existing product technologies and the failure of established firms. *Administrative Science Quarterly* 35 (1): 9–30.
Heras-Saizarbitoria, I., and O. Boiral. 2013. Research agenda on management system standards. *International Journal of Management Reviews* 15: 47–65.
Hill, C. A. 2004. Regulating the rating agencies. *Washington University Law Quarterly* 82(1): 43–94.
Hoffman, A. J. 2001. *From Heresy to Dogma: An Institutional History of Corporate Environmentalism*. Stanford, CA: Stanford University Press.
Hoffman, V. B., and J. M. Patton. 1997. Accountability, the dilution effect, and conservatism in auditors' fraud judgments. *Journal of Accounting Research* 35(2): 227–237.
Huber, G.R. and D.J. Power. 1985. Retrospective reports of strategic-level managers: Guidelines for increasing their accuracy. *Strategic Management Journal* 6: 171–180.
Jacob, B. A. 2002. Where the boys aren't: Non-cognitive skills, returns to school, and the gender gap in higher education. *Economics and Education Review* 21: 589–598.
Jenkins, M., G. Esquivel, and F. Larrían. 2001. Export processing zones in Central America. In Felipe Larraín (ed.), *Economic development in Central America, vol. I: Growth and internationalization:* 197–150. Boston: John F. Kennedy School of Government, Harvard University.
Jones, B. D. 2001. *Politics and the Architecture of Choice: Bounded Rationality and Governance*. Chicago: University of Chicago Press.
Jonnergård, K., A. Stafsudd, and U. Elg. 2010. Performance evaluations as gender barriers in professional organizations: A study of auditing firms. *Gender, Work & Organization* 17(6): 721–747.
Joshi, A., and H. Roh. 2009. The role of context in work team diversity research: A meta-analytic review. *Academy of Management Journal* 52(3): 599–627.
Kamatali, J-M. 2012. The new guiding principles on human rights' contribution in ending the devising debate over human rights responsibilities of companies: Is it time for an ICJ advisory opinion? *Cardozo Journal of International and Comparative Law* 20(2): 437–463.
Kanter, R. M. 1977. *Men and Women of the Corporation*. New York: Basic Books.
Keiser, L. R., and J. Soss. 1998. With good cause: Bureaucratic discretion and the politics of child support enforcement. *American Journal of Political Science* 42(4): 1133–1156.
Killgore, W. D. S., and D. W. Cupp. 2002. Mood and sex of participant in perception of happy faces. *Perceptual and Motor Skills* 95: 279–288.
King, A. A., and M. W. Toffel. 2009. Self-regulatory institutions for solving environmental problems: Perspectives and contributions from the management literature. In M. Delmas and O. Young (eds.), *Governance for the Environment: New Perspectives:* 98–115. New York: Cambridge University Press.
King, A., and M. Lenox. 2001. Lean and green? An empirical examination of the relationship between lean production and environmental performance. *Production and Operations Management* 10(3): 244–257
King, A., M. Lenox, and A. Terlaak. 2005. The strategic use of decentralized institutions:

Exploring certification with the ISO 14001 management standard. *Academy of Management Journal* 48(6): 1091–1106.
Kocer, G. R., and L. Fransen. 2009. Codes of conduct and the promise of a change of climate in worker organization. *European Journal of Industrial Relations* 15(3): 237–256.
Kumar, J.A., and A. Chakrabarti. 2012. Bounded awareness and tacit knowledge: Revisiting Challenger disaster. *Journal of Knowledge Management* 16(6): 934–949.
Kuncoro, M. 2011. *Recent Development of the Indonesian Apparel Industry*. Yogyakarta, Indonesia: Gadjah Mada University.
Larson, M. S. 1977. *The Rise of Professionalism: A Sociological Analysis*. Berkeley, CA: University of California Press.
Lemley, M. A., and B. Sampat. 2012. Examiner characteristics and patent office outcomes. *Review of Economics and Statistics* 94(3): 817–827.
Levine, D. I., and M. W. Toffel. 2010. Quality management and job quality: How the ISO 9001 standard for quality management systems affects employees and employers. *Management Science* 56(6): 978–996.
Lipsky, M. 1980/2010. *Street-Level Bureaucracy: Dilemmas of the Individual in Public Services*. New York: Russell Sage Foundation.
Locke, R. M., and M. Romis. 2007. Improving work conditions in a global supply chain. *MIT Sloan Management Review* 48(2): 54–62.
Locke, R. M., B. A. Rising, and T. Pal. 2012. Complements or substitutes? Private codes, state regulation and the improvement of labor standards in global supply chains. Research paper no. 2012-2, MIT Political Science Department.
Locke, R. M., F. Qin, and A. Brause. 2007. Does monitoring improve labor standards? Lessons from Nike. *Industrial and Labor Relations Review* 61(1): 3–31.
Lount, R. B., Jr., and K. W. Phillips. 2007. Working harder with the out-group: The impact of social category diversity on motivation gains. *Organizational Behavior and Human Decision Processes* 103: 214–224.
Loyd, D. L., C. S. Wang, K. W. Phillips, and R. B. Lount, Jr. 2013. Social category diversity promotes premeeting elaboration: The role of relationship focus. *Organization Science* 24(3): 757–772.
Macher, J. T., J. W. Mayo, and J. A. Nickerson. 2011. Regulator heterogeneity and endogenous efforts to close the information asymmetry gap. *Journal of Law and Economics* 54(1): 25–54.
Manns, J. 2013. Downgrading rating agency reform. *George Washington Law Review* 81: 749–812.
May, P. J., and S. Winter. 2000. Reconsidering styles of regulatory enforcement: Patterns in Danish agro-environmental inspection. *Law and Policy* 22(2): 143–173.
Maynard-Moody, S., and M. Musheno. 2003. *Cops, Teachers, Counselors: Narratives of Street-Level Judgment*. Ann Arbor, MI: University of Michigan Press.
McBarnet, D. 2007. Corporate social responsibility beyond law, through law, for law: The new corporate accountability. In D. McBarnet, A. Voiculescu, and T. Campbell (eds.), *The New Corporate Accountability: Corporate Social Responsibility and the Law:* 9–56. Cambridge, UK: Cambridge University Press.
McBarnet, D., and M. Kurkchiyan. 2007. Corporate social responsibility through contractual control? Global supply chains and 'other-regulation.' In D. McBarnet, A. Voiculescu, and T. Campbell (eds.), *The New Corporate Accountability: Corporate Social Responsibility and the Law:* 59–92. Cambridge, UK: Cambridge University Press.

Meidinger, Errol E. 2002. The new environmental law: Forest certification. *Buffalo Environmental Law Journal* 10: 213–303.
Montiel I., B.W. Husted, and P. Christmann. 2012. Using private management standard certification to reduce information asymmetries in corrupt environments. *Strategic Management Journal* 33: 1103–1113.
Moore, D. A., P. E. Tetlock, L. Tanlu, and M. H. Bazerman. 2006. Conflicts of interest and the case of auditor independence: Moral seduction and strategic issue cycling. *Academy of Management Review* 31(1): 10–29.
Mueller, F., C. Carter, and A. Ross-Smith. 2011. Making sense of career in a Big Four accounting firm. *Current Sociology* 59(4): 551–567.
O'Rourke, D. 2002. Monitoring the monitors: A critique of PricewaterhouseCooper's labor monitoring. In R. Jenkins, R. Pearson and G. Seyfang (eds.), *Corporate Responsibility and Ethical Trade: Codes of Conduct in the Global Economy*: 196–207. London: Earthscan
O'Rourke, D. 2003. Outsourcing regulation: Analyzing nongovernmental systems of labor standards and monitoring. *Policy Studies Journal* 31(1): 1–29.
Oberfield, Z. W. 2010. Rule following and discretion at government's frontlines: Continuity and change during organization socialization. *Journal of Public Administration Research and Theory* 20(4): 735–755.
Oh, P. B. 2004. Gatekeeping. *Journal of Corporation Law* 29(4): 735–800.
Parker, C. 1999. Compliance professionalism and regulatory community: The Australian trade practices regime. *Journal of Law and Society* 26(2): 215–239.
Parmigiani, A. 2007. Why do firms both make and buy? An investigation of concurrent sourcing. *Strategic Management Journal* 28: 285–311.
Partnoy, F. 2004. Strict liability for gatekeepers: A reply to Professor Coffee. *Boston University Law Review* 84: 365–375.
Partnoy, F. 2006. How and why credit rating agencies are not like other gatekeepers. In Y. Fuchita and R. E. Litan (eds.), *Financial Gatekeepers: Can They Protect Investors?*: 59–99. Baltimore, MD: Brookings Institution Press.
Phillips, K. W., M. Duguid, M., Thomas-Hunt, and J. Uparna. 2012. Diversity as knowledge exchange: The roles of information processing, expertise, and status. In Q. M. Roberson (ed.), *The Oxford handbook of diversity and work:* 157–178. New York: Oxford University Press.
Pierce, L., and M. W. Toffel. 2013. The role of organizational scope and governance in strengthening private monitoring. *Organization Science* 24(5): 1558–1584.
Piore, M. J. 2005. Rethinking international labor standards. In W. Milberg (ed.), *Labor and the Globalization of Production: Causes and Consequences*: 249–265. New York: Macmillan.
Piore, M. J., and A. Schrank. 2008. Toward managed flexibility: The revival of labour inspection in the Latin world. *International Labour Review* 147(1): 1–23.
Portillo, S. 2012. The paradox of rules: Rules as resources and constraints. *Administration and Society* 44(1): 87–108.
Portillo, S., and L. DeHart-Davis. 2009. Gender and organizational rule abidance. *Public Administration Review* 69: 339–347.
Potoski, M., and A. Prakash. 2004. Regulatory convergence in nongovernmental regimes: Cross-national adoption of ISO 14001 certification. *Journal of Politics* 66: 885–905.
Potoski, M., and A. Prakash. 2005a. Covenants with weak swords: ISO 14001 and facilities' environmental performance. *Journal of Policy Analysis Management* 24: 745–769.
Potoski, M., and A. Prakash. 2005b. Green clubs and voluntary governance: ISO 14001 and

firms' regulatory compliance. *American Journal of Political Science* 49(2): 235–248.
Power, D. and M. Terziovski. 2007. Quality audit roles and skills: Perceptions of non-financial auditors and their clients. *Journal of Operations Management* 25(1): 126–147.
Power, M. 1997. *The Audit Society: Rituals of Verification*. Oxford: Oxford University Press.
Prakash, A., and M. Potoski. 2006. Racing to the bottom? Trade, environmental governance, and ISO 14001. *American Journal of Political Science* 50(2): 350–364.
Reitzig M. and S. Wager. 2010. The hidden costs of outsourcing: Evidence from patent data. *Strategic Management Journal* 31: 1183–1201.
Rodríguez-Garavito, C. 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.
Ruggie, J. 2008. *Protect, Respect and Remedy: A Framework for Business and Human Rights. Report of the Special Representative of the Secretary-General on the Issue of Human Rights and Transnational Corporations and Other Business Enterprises*. United Nations, Human Rights Council, Geneva, Switzerland. http://www.reports-and-materials.org/Ruggie-report-7-Apr-2008.pdf.
Sandfort, J. R. 2000. Moving beyond discretion and outcomes: Examining public management from the front lines of the welfare system. *Journal of Public Administration Research and Theory* 10(4): 729–756.
Schrank, A. 2009. Professionalization and probity in a patrimonial state: Labor inspectors in the Dominican Republic. *Latin American Politics and Society* 51(2): 91–115.
Schwarcz, S. L. 2002. Private ordering of public markets: The rating agency paradox. *University of Illinois Law Review* 2002(1): 1–27.
Scott, C. 2012. Non-judicial enforcement of transnational private regulation. In F. Cafaggi (ed.), *Enforcement of Transnational Regulation: Ensuring Compliance in a Global World:* 147–164. Cheltenham, UK: Edward Elgar.
Scott, R. W. 1966. Professionals in bureaucracies—areas of conflict. In H. M. Vollmer and D. L. Mills (eds.), *Professionalization:* 265–275. Englewood Cliffs, NJ: Prentice-Hall.
Shamir, R. 2005. Corporate social responsibility: A case of hegemony and counter-hegemony. In B. de Sousa Santos & C. A. Rodríguez-Garavito (Eds.), *Law and Globalization from Below:* 92–117. Cambridge: Cambridge University Press.
Short, J.L, and M.W. Toffel. 2010. Making self-regulation more than merely symbolic: The critical role of the legal environment. *Administrative Science Quarterly* 55(3): 361–396.
Simon, H. A. 1947. *Administrative Behavior: A Study of Decision-Making Processes in Administrative Organizations*. New York: Free Press.
Spence, J. T., and C. E. Buckner. 2000. Instrumental and expressive traits, trait stereotypes, and sexist attitudes: What do they signify? *Psychology of Women Quarterly* 24: 44–62.
Stonewater, B. B., S. A. Eveslage, and M. R. Dingerson. 1990. Gender differences in career helping relationships. *Career Development Quarterly* 39: 72–85.
Surroca J., J.A. Tribo, and S. Waddock. 2010. Corporate responsibility and financial performance: The role of intangible resources. *Strategic Management Journal* 31: 463–490.
Teitelbaum, J. C. 2006. Age and tenure of the justices and productivity of the Supreme Court: Are term limits necessary? *Florida State University Law Review* 34: 161–182.
Terlaak, A., and A. A. King. 2006. The effect of certification with the ISO 9000 Quality Management Standard: A signaling approach. *Journal of Economics Behavior and Organization* 60(4): 579–602.

Tetlock, P. E. 1983. Accountability and complexity of thought. *Journal of Personality and Social Psychology* 45(1): 74–83.
Thayer, J. F., and B. H. Johnsen. 2000. Sex differences in judgment of facial affect: A multivariate analysis of recognition errors. *Scandinavian Journal of Psychology* 41(3): 243–246.
Thompson, J. D. 1967. *Organizations in Action*. New York: McGraw-Hill.
Thompson, V. A. 1977. *Modern Organization*, 2nd ed. New York: Alfred A. Knopf.
Toffel, M. W., J. L. Short, and M. Ouellet. 2013. Codes in context: How states, markets, and civil society shape adherence to global labor standards. Working paper 13–045, Harvard Business School.
Turner, C. W. 2001. Accountability demands and the auditor's evidence search strategy: The influence of reviewer preferences and the nature of the response (belief vs. action). *Journal of Accounting Research* 39(3): 683–706.
U.S. Public Company Accounting Oversight Board. 2011. *Concept Release on Auditor Independence and Audit Firm Rotation*. Release No. 2011-006 (August 16).
US SIF. 2012. *Report on Sustainable and Responsible Investing Trends in the United States: 2012*. Washington, D.C.: The Forum for Sustainable and Responsible Investment (US SIF).
von Krogh, G., J. Roos, and K. Slocum. 1994. An essay on corporate epistemology. *Strategic Management Journal* 15(1): 53–71.
Waddock, S.A., and S.B. Graves. 1997. The corporate social performance-financial performance link. *Strategic Management Journal* 18(4): 303–319.
Walsh, D., and S. Greenhouse. 2012. Certified safe, a factory in Karachi still quickly burned. *New York Times*, December 8: A1.
Weber, B., M. Wittchen, and G. Hertel. 2009. Gendered ways to motivation gains in groups. *Sex Roles* 60: 731–744.
Weber, M. 1947. *The Theory of Social and Economic Organization*. New York: Free Press.
Williamson, A., J. H. Rogerson, and A. D. Vella. 1996. Quality system auditors' attitudes and methods: A survey. *International Journal of Quality Reliability and Management* 13(8): 39–52.
Wood, S. 2004. The role of the International Organization for Standardization (ISO) in governing environmental conflict and corporate social responsibility in developing countries: Questions for research. In B. L. Toro (ed.), *Propriedad, Conflicto y Medio Ambiente*: 15–56. Bogotá: Universidad del Rosario.
World Bank. 2013. Worldwide governance indicators: Frequently asked questions. http://info.worldbank.org/governance/wgi/faq.htm, accessed April 2013.
Yardley, J. 2012. Recalling fire's horror and exposing global brands' safety gap. *New York Times*, December 7: A1.
Yin, H., and P. J. Schmeidler. 2009. Why do standardized ISO 14001 environmental management systems lead to heterogeneous environmental outcomes? *Business Strategy and the Environment* 18: 469–486.
Zandvliet, R. 2011. Corporate social responsibility reporting in the European Union: Towards a more univocal framework. *Columbia Journal of European Law, On Line* 18: 38.

Table 1. Industry composition

Industry	Audits		Factories	
	Number	Percent	Number	Percent
Accessories	1,740	10	579	10
Building materials	260	2	84	1
Chemicals and plastics	97	1	42	1
Electronics	590	4	184	3
Food, agriculture, beverage	138	1	58	1
Footwear	356	2	122	2
Furniture	383	2	123	2
Garments	6,188	37	2,113	36
Metal products	156	1	51	1
Paper, printing, publishing	183	1	63	1
Services	50	0	19	0
Toys	463	3	150	3
Other/unknown	6,191	37	2,231	38
Total	16,795	100	5,819	100

Table 2. Summary statistics

Variable	**Mean**	**SD**	**Min**	**Max**
Number of violations	6.49	5.61	0	75
Previous auditor	0.15	0.36	0	1
Maximum tenure	5.39	2.03	1	15
Average tenure	4.86	1.85	0.5	15
Graduate education	0.13	0.34	0	1
Auditing skills training	2.25	1.74	0	12
All-male audit team	0.33	0.47	0	1
All-female audit team	0.50	0.50	0	1
Mixed-gender audit team	0.17	0.37	0	1
Certification training	0.50	0.42	0	1
Brand training	0.59	0.43	0	1
Average age	30.12	4.47	22.5	59
Maximum age	30.62	4.66	25	59
Third-party protocol	0.04	0.19	0	1
Unannounced audit	0.22	0.41	0	1
Audit paid for by factory or agent	0.56	0.50	0	1
Audit paid for by brand	0.44	0.50	0	1
Re-audit	0.36	0.48	0	1
Number of auditors	1.79	0.58	1	5
Audit sequence	2.96	2.25	1	21
Per-capita GDP (log)	7.77	0.98	5.61	10.68
Regulatory quality	-0.04	0.54	-1.64	1.99
Press freedom	0.33	0.27	0.12	1.00

Note: N=16,795 audits except N=15,812 for *audit paid for by factory or agent* and *audit paid for by brand*, N=11,337 for *average age* and *maximum age* and N=16,676 for *press freedom*.

Table 3. Pairwise Correlations

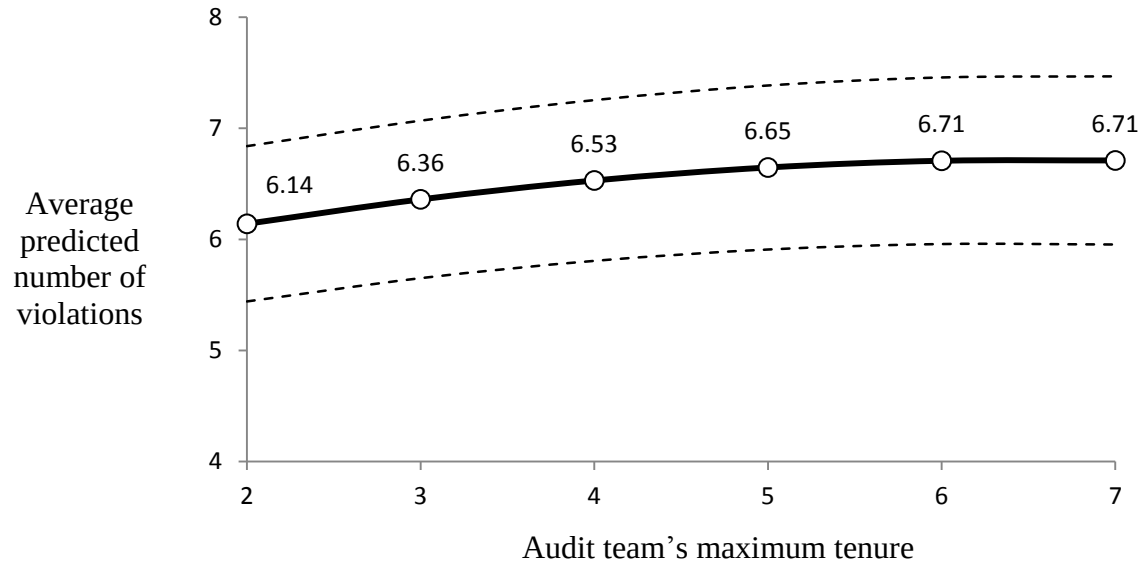
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)
(1) Number of violations	1.00																					
(2) Previous auditor	-0.13	1.00																				
(3) Maximum tenure	-0.01	0.03	1.00																			
(4) Average tenure	-0.03	0.03	0.92	1.00																		
(5) Graduate education	-0.07	0.09	-0.04	-0.04	1.00																	
(6) Auditing skills training	-0.03	0.05	-0.20	-0.30	-0.04	1.00																
(7) All-male audit team	-0.07	0.07	-0.06	-0.01	0.07	-0.07	1.00															
(8) All-female audit team	0.05	-0.09	0.03	0.07	-0.14	0.01	-0.70	1.00														
(9) Mixed-gender audit team	0.02	0.04	0.03	-0.09	0.09	0.07	-0.32	-0.45	1.00													
(10) Certification training	0.01	0.04	0.09	0.12	-0.04	0.47	0.04	0.03	-0.09	1.00												
(11) Brand training	-0.02	0.06	-0.15	-0.16	-0.09	0.59	0.02	-0.01	-0.02	0.45	1.00											
(12) Average age	-0.10	0.16	0.43	0.49	0.18	-0.17	0.22	-0.22	0.02	0.08	-0.12	1.00										
(13) Maximum age	-0.08	0.16	0.43	0.45	0.19	-0.12	0.19	-0.23	0.07	0.07	-0.12	0.97	1.00									
(14) Third-party protocol	0.12	0.01	-0.02	-0.03	0.05	0.00	0.04	-0.05	0.01	0.03	0.01	0.00	0.00	1.00								
(15) Unannounced audit	0.00	0.05	0.11	0.11	0.02	-0.01	0.02	-0.02	-0.01	-0.03	-0.04	0.06	0.05	-0.09	1.00							
(16) Audit paid for by factory or agent	0.02	-0.07	0.00	0.01	-0.09	-0.07	-0.05	0.04	0.01	0.01	-0.04	-0.03	-0.02	-0.14	-0.20	1.00						
(17) Audit paid for by brand	-0.02	0.07	0.00	-0.01	0.09	0.07	0.05	-0.04	-0.01	-0.01	0.04	0.03	0.02	0.14	0.20	-1.00	1.00					
(18) Re-audit	-0.12	0.12	-0.04	-0.05	-0.05	0.01	-0.02	0.02	-0.01	-0.02	-0.01	-0.09	-0.08	-0.02	0.07	0.05	-0.05	1.00				
(19) Number of auditors	0.13	-0.05	-0.06	-0.27	0.01	0.04	-0.23	-0.05	0.36	-0.27	-0.18	-0.29	-0.21	0.04	-0.03	0.04	-0.04	0.05	1.00			
(20) Audit sequence	-0.28	0.15	-0.14	-0.20	-0.03	0.31	-0.04	-0.01	0.06	0.09	0.20	-0.11	-0.09	-0.07	0.02	0.01	-0.01	0.12	0.02	1.00		
(21) Per-capita GDP (log)	-0.18	0.08	0.22	0.20	-0.05	0.21	-0.04	0.06	-0.03	0.09	0.14	-0.06	-0.08	-0.07	0.12	-0.16	0.16	-0.09	-0.13	0.14	1.00	
(22) Regulatory quality	-0.19	0.10	0.22	0.21	0.01	0.14	-0.02	0.04	-0.02	0.07	0.09	0.01	-0.02	-0.06	0.13	-0.18	0.18	-0.11	-0.15	0.11	0.95	1.00
(23) Press freedom	-0.27	0.22	0.22	0.25	0.34	-0.05	0.15	-0.14	0.00	-0.02	-0.02	0.46	0.42	0.00	0.12	-0.24	0.24	-0.14	-0.28	0.02	0.54	0.65

Table 4. Regression results

	(1)		(2)	(3)	(4)	(5)
	Coef.	Average marginal effects	Coef.	Coef.	Coef.	Coef.
H1 Previous auditor	-0.043*	-0.28	-0.039+	-0.028	-0.044*	-0.027
	[0.020]		[0.020]	[0.026]	[0.021]	[0.025]
H2 Maximum tenure	0.065**	0.12	0.068**	0.078**	0.069**	0.084**
	[0.014]		[0.014]	[0.016]	[0.014]	[0.017]
H2 Maximum tenure, squared	-0.004**		-0.004**	-0.004**	-0.005**	-0.005**
	[0.001]		[0.001]	[0.001]	[0.001]	[0.001]
H3 Graduate education	0.027	0.18	0.030	-0.004	0.021	0.045
	[0.024]		[0.024]	[0.029]	[0.026]	[0.039]
H3 Auditing skills training	0.021**	0.14	0.022**	0.013	0.022**	0.012
	[0.007]		[0.007]	[0.009]	[0.007]	[0.009]
H4 All-female audit team	0.054**	0.35	0.055**	0.048*	0.053**	0.052*
	[0.015]		[0.015]	[0.019]	[0.016]	[0.021]
H5 Mixed-gender audit team	0.067**	0.43	0.068**	0.049*	0.069**	0.067*
	[0.021]		[0.021]	[0.024]	[0.021]	[0.028]
Certification training	-0.021	-0.14	-0.024	-0.027	-0.031	-0.010
	[0.021]		[0.021]	[0.027]	[0.021]	[0.029]
Brand training	-0.014	-0.09	-0.012	0.008	-0.007	0.001
	[0.021]		[0.022]	[0.026]	[0.022]	[0.030]
Average age	-0.025	-0.04	-0.026	-0.015	-0.023	-0.041
	[0.019]		[0.020]	[0.028]	[0.019]	[0.027]
Average age, squared	0.000		0.000	0.000	0.000	0.001
	[0.000]		[0.000]	[0.000]	[0.000]	[0.000]
Third-party protocol	-0.080	-0.52	-0.088	-0.148*		-0.210*
	[0.058]		[0.062]	[0.070]		[0.101]
Unannounced audit	0.029	0.19	0.029	0.030	0.031	0.075**
	[0.020]		[0.020]	[0.025]	[0.020]	[0.027]
Audit paid for by factory or agent	-0.084**	-0.55	-0.083**	-0.068*	-0.064*	-0.099**
	[0.026]		[0.027]	[0.032]	[0.028]	[0.034]
Re-audit	-0.348**	-2.26	-0.351**	-0.353**	-0.358**	-0.345**
	[0.016]		[0.016]	[0.019]	[0.017]	[0.020]
Per-capita GDP (log)	-0.623*	-4.04	-0.551*	-0.749	-0.714**	-0.210
	[0.262]		[0.264]	[0.473]	[0.267]	[0.389]
Regulatory quality	-0.180	-1.17	-0.169	-0.385	-0.158	-0.621**
	[0.150]		[0.150]	[0.298]	[0.153]	[0.231]
Press freedom	-0.510*	-3.31	-0.531*	-1.059*	-0.402+	-0.879**
	[0.224]		[0.224]	[0.476]	[0.239]	[0.339]
Observations (audits)	16,795		16,585	10,648	16,044	9,266
Factories	5,819		5,748	3,810	5,523	3,082

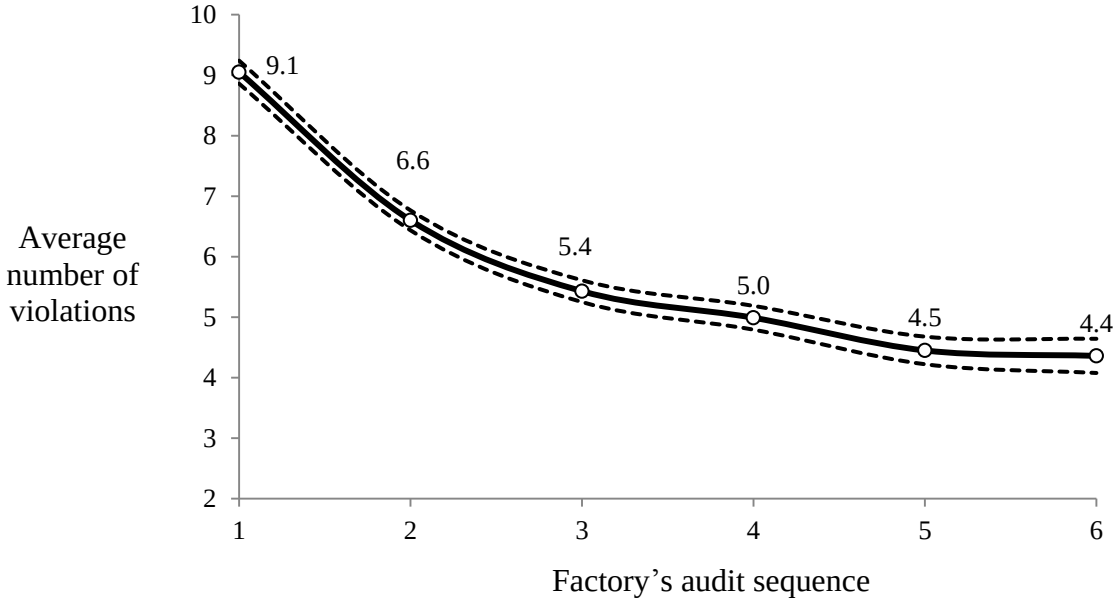
Standard errors clustered by audited factory; ** $p < 0.01$, * $p < 0.05$, + $p < 0.10$. All models also include fixed effects for the audited establishment, audit year (2005 through 2009), client country, number of auditors (2 through 5 or more), and the factory's audit sequence (2^{nd} through 6^{th} or more). All models include three dummy variables to indicate instances in which the following variables were missing data and thus recoded to 0: *average age* and *maximum age* (5,458 audits), *audit paid for by factory or agent* and *audit paid for by brand* (983 audits), and *press freedom* (119 audits). Model 1 is the primary model estimated on the entire sample. Model 2 excludes audits conducted for clients whose audit teams were always all-female. Model 3 includes only audits conducted by at least two auditors. Model 4 excludes audits conducted according to a third-party protocol. Model 5 excludes factories' first inspection during the sample period.

Figure 1. Effect of the audit team's maximum tenure on average predicted violations per factory



Note: The figure depicts average predicted number of violations from the fixed-effects Poisson model estimated in Column 1 of Table 4, spanning the 5th to 95th percentiles of audit tenure. Dashed lines represent the 95-percent confidence interval.

Figure 2. The average number of violations per audit declines in factories' successive audits



Note: The figure depicts sample averages, with dashed lines representing 95-percent confidence intervals calculated as the sample mean ± two times the standard error of the mean.