

2014

Monitoring the Monitors How Social Factors Influence Supply Chain Auditors

Jodi L. Short

UC Hastings College of the Law, shortj@uchastings.edu

Michael W. Toffel

Andrea R. Huguill

Follow this and additional works at: http://repository.uchastings.edu/faculty_scholarship

Recommended Citation

Jodi L. Short, Michael W. Toffel, and Andrea R. Huguill, *Monitoring the Monitors How Social Factors Influence Supply Chain Auditors*, *Strategic Management Journal* ____ (2014).

Available at: http://repository.uchastings.edu/faculty_scholarship/1194

This Article is brought to you for free and open access by UC Hastings Scholarship Repository. It has been accepted for inclusion in Faculty Scholarship by an authorized administrator of UC Hastings Scholarship Repository. For more information, please contact marcus@uchastings.edu.

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

September 4, 2014

Outsourcing firms increasingly rely on social auditors to provide strategic information about the conduct of their suppliers to manage the reputational risks that can arise from dangerous, illegal, and unethical behavior at supply chain factories. But little is known about what influences auditors' ability to identify and report poor supplier conduct. We find evidence that private supply chain auditors' reporting practices are shaped by several social factors including their experience, gender, and professional training; their ongoing relationships with suppliers; and the gender diversity of their audit teams. By providing the first comprehensive and systematic findings on supply chain auditing practices, our study suggests strategies companies can pursue to develop more credible monitoring regimes to reduce information asymmetries between themselves and their suppliers.

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

Reputation is a key strategic concern for modern firms (Roberts and Dowling, 2002) and environmental, social, and governance (ESG) performance have become increasingly important dimensions of the firm's reputation. Poor ESG practices may make firms targets for activists, harming their reputations as well as their bottom lines (King and Soule, 2007; Vasi and King, 2012), and firms are increasingly being held accountable by financial analysts and investors for their ESG performance (Ioannou, 2014). Furthermore, firms that mislead stakeholders about their ESG performance through inaccurate disclosures may find their reputations tarnished

[†] 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, Jeffrey Macher, Andrew Marder, Shane Greenstein, Morris Ratner, Bill Simpson, Veronica Villena, 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.

(Kayser, Maxwell, and Toffel, 2014).

As firms continue to outsource much of their production globally, their reputations have come to depend not solely on their own practices but also on those of the companies in their extended supply chains (Heide, Kumar, and Wathne, 2014). For instance, in the wake of the Rana Plaza building collapse in Bangladesh that killed 1,100 factory workers, the bulk of media, consumer, and activist scrutiny focused on the global retailers that sourced from suppliers using the building (Greenhouse, 2013a). Many of these global retailers were under such intense reputational pressure that they agreed to adopt a legally binding accord requiring them to spend hundreds of millions of dollars to fund fire-safety and structural improvements in the factories of their Bangladeshi suppliers (Greenhouse, 2013b).

Largely because of the risk of negative reputational spillovers, supply chain conditions are listed as a top stakeholder concern in global indices (such as Innovest and the Global Reporting Initiative) and in recent strategy research (Crilly and Sloan, 2012). Consumers, investors, and activists are increasingly holding firms accountable for the poor social practices of their suppliers (O’Callaghan, 2007) and research suggests that non-governmental organizations can impose substantial reputational penalties on firms whose suppliers violate global norms (Fung, O’Rourke, and Sabel, 2001; Vogel, 2010). Reputational risk is particularly high for firms that have marketed their ethical sourcing practices to consumers if facts come to light that contradict these claims. Studies show that firms risk being punished when stakeholders suspect them of “greenwashing” or “bluwashing” to mislead stakeholders by emphasizing positive attributes while neglecting to mention negative ones (Bowen, 2014; Ioannou, 2014; Kayser, Maxwell, and Toffel, 2014).

To manage such risk, many firms require their suppliers to meet globally recognized

standards in areas including environmental sustainability, working conditions, and human rights. These requirements are typically imposed through code-of-conduct provisions in supply contracts or through voluntary certification schemes such as SA8000, Fairtrade, or the Forest Stewardship Council. Suppliers' adherence to these standards is often enforced through private social monitoring by supply chain auditors, also referred to as "social auditors" (Montiel, Husted, and Christmann, 2012).

Social monitoring to manage supply chain risk has become a key governance mechanism providing firms with information with which to make strategic outsourcing decisions. However, supply chain auditors have been criticized for failing to provide firms with complete and accurate information about the conditions in their supply chains (e.g., Esbenshade, 2004; Heras-Saizarbitoria and Boiral, 2013; O'Rourke, 2002). When auditors fail to accurately assess suppliers' adherence to standards, they undermine firms' ability to make fully informed outsourcing decisions and they subject firms to the risk of catastrophic reputational consequences. For instance, in 2012, just weeks after social auditors certified that a factory in Pakistan met the SA8000 working conditions standard created by a respected nonprofit, a fire there killed hundreds of workers, some of whom were trapped by locked emergency exits and barred windows—clear violations of that standard (Walsh and Greenhouse, 2012).

Despite the importance of supply chain auditors to firms' reputation management, little is known about how they do their jobs. Much of the literature on private supply chain standards brackets the question of auditor performance or implicitly assumes that auditors provide reasonably objective assessments of on-the-ground conditions. To the extent that scholars address auditor performance, they have typically argued anecdotally that supply chain auditors are biased in favor of their paying clients (Esbenshade, 2004; Heras-Saizarbitoria and Boiral,

2013; O'Rourke, 2002). Recent qualitative research reveals that auditors often deny their vulnerability to conflicts of interest and socially construct a professional identity that emphasizes their independence (Dogui, Boiral, and Gendron, 2013). To our knowledge, however, no empirical research has rigorously investigated what factors shape supply chain auditors' assessments of supplier adherence to standards.

We seek to fill that gap. Grounding our work in the strategic management literature and taking insights from the literatures on regulatory compliance and on financial auditing, we theorize and investigate individual and team characteristics that affect how social auditors do their job as they assess and record supply chain factories' violations of private standards. We argue that auditors are not merely objective transmitters of supply chain conditions but instead that social relationships, institutions, and individual identities shape the information they transmit to their corporate clients.

We test our hypotheses in the context of social auditing for compliance with labor standards contained in corporate supplier codes of conduct. This form of private supply chain regulation has been adopted by thousands of prominent multinational corporations (MNCs), including all U.S. *Fortune* 500 companies (McBarnet, 2007). We exploit a novel dataset drawn from thousands of audits for code-of-conduct compliance in over 66 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 such as 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.

LITERATURE REVIEW

Decisions about outsourcing and the management of supply chains have become increasingly critical strategic concerns for firms (Alcacer and Oxley, 2014; Hult *et al.*, 2007; Parmigiani, 2007; Reitzig and Wagner, 2010). Current concerns go beyond the paradigmatic “make or buy” dilemma (Williamson, 1975) to raise complex strategic questions about *which suppliers to buy from*. Choosing the correct supplier involves considering not only cost and quality, but also a variety of other concerns such as flexibility, reliability, and management attitude that can likewise affect firm performance. On the one hand, there is growing evidence that choosing suppliers wisely can enhance a firm’s value by, for instance, improving its financial performance (Doig *et al.*, 2001) or providing opportunities to develop knowledge (Alcacer and Oxley, 2014; Hult *et al.*, 2004, 2007). On the other hand, poor supplier choices can subject firms to serious costs, including remediation expenses, legal liability, and reputation damage (O’Callaghan, 2007).

To reduce reputational risks and protect brand value, firms have invested significant resources to monitor their suppliers’ behavior (Blair *et al.*, 2008). As a condition of doing business, most multinational firms require their suppliers to meet globally recognized standards in areas including environmental sustainability, working conditions, and human rights (McBarnet, 2007) and many employ supply chain auditors to monitor suppliers’ adherence to these standards (Montiel, Husted, and Christmann, 2012). The transaction cost economics (TCE) literature has long posited that monitoring is a key mechanism for reducing transaction costs that arise when contractual partners opportunistically exploit information asymmetries (Eisenhardt, 1985; Heide, Kumar, and Wathne, 2014). “[M]onitoring offers control by reducing information asymmetry between exchange parties” (Heide, Wathne, and Rokkan, 2007: 426). Specifically,

the information gathered by monitors provides firms “with strategically valuable knowledge about how [their exchange partners] operate and the perception of their operations among stakeholders” (Park, 2014: 28). Ideally, such information should give a firm an accurate picture of the risks and benefits posed by different suppliers, allowing it to make crucial strategic decisions about which suppliers to work with and where.

While the TCE literature highlights the important governance function of monitoring, it contains little empirical research on monitoring and has not conceptualized the distinct role and contours of monitoring in global supply chains. First, the TCE literature has focused on “ex ante safeguards to deter ex post opportunism” (Williamson, 2008), including contractual provisions (Argyres, Bercovitz, and Mayer, 2007; Barthélemy and Quélin, 2006; Kalnins and Mayer, 2004), technological capabilities (Mayer and Salomon, 2006), and shared knowledge base (Garicano and Hubbard, 2009; Heide, Kumar, and Wathne, 2014; Puranam *et al.*, 2013). This focus has obscured the importance of ongoing, *ex post* monitoring to contain the costs of negative reputational spillovers from contractual partners’ poor social practices.

Second, the TCE literature has not appreciated the fact that the monitoring of outsourced production in supply chains is, itself, often outsourced to third-party auditors. Although some firms have extensive in-house monitoring programs, most rely on private social auditors, certification organizations, or other independent assurance services to monitor their suppliers and provide them with strategic information about exchange partners’ activities. Outsourcing of the firm’s traditional monitoring function has created second-order agency problems associated with monitoring the monitors. These problems have not been adequately addressed in scholarship or in practice. Firms tend to assume that the information they receive from the monitors they retain accurately reflects conditions on the ground in their supply chains, but critics have suggested that

this is not always the case (e.g., Esbenshade, 2004; Heras-Saizarbitoria and Boiral, 2013; O'Rourke, 2002; Pierce and Toffel, 2013). To date, however, neither TCE research nor research on supply chain auditing has provided guidance on how to enhance the reliability of the information firms receive about their suppliers through monitoring.

There is a substantial literature on the implementation of private standards governing supply chain practices, including how these standards are developed (Meidinger, 2002; Wood, 2004), why they are adopted (Bartley, 2007, 2010; Boiral, 2007; Christmann and Taylor, 2001, 2006; Delmas, 2002; Delmas and Montiel, 2008; Delmas and Toffel, 2008; Guler, Guillén, and MacPherson, 2002; Hoffman, 2001; King, Lenox, and Terlaak, 2005; Potoski and Prakash, 2004; Prakash and Potoski, 2006; Terlaak and King, 2006), and their financial, operational, and compliance outcomes (Corbett, Montes-Sancho, and Kirsch, 2005; Delmas, 2001; Esbenshade, 2004; King and Lenox, 2001; Kocer and Fransen, 2009; Levine and Toffel, 2010; Locke, Rising, and Pal, 2012; Locke and Romis, 2007; Potoski and Prakash, 2005a, 2005b; Rodríguez-Garavito, 2005; Short and Toffel, 2010; Terlaak and King, 2006; Toffel and Short, 2011; Yin and Schmeidler, 2009). However, while the literature examining supply chain standards assigns private auditors a leading role, it provides little insight into how they play it (Heras-Saizarbitoria and Boiral, 2013).

In fact, most research on supply chain standards assumes that private auditors hired to monitor supplier compliance document and report violations in a way that is reasonably objective and that depends principally on the scope and content of the standards contained in the audit protocol. Such assumptions, however, are contrary to studies of public regulatory implementation documenting significant heterogeneity in the way government monitors apply the rules they are charged with enforcing (Black, 1997; Feinstein, 1989, 1990; Hawkins, 1984;

Lemley and Sampat, 2012; Macher, Mayo, and Nickerson, 2011) and to experimental studies in the social psychology literature documenting the influence of cognitive biases on the performance of financial auditors (e.g., Asare, Trompeter, and Wright, 2000; Glover, 1997; Hoffman and Patton, 1997; Tetlock, 1983; Turner, 2001).

Some empirical studies have investigated influences on the performance of private-sector auditors, but have focused almost exclusively on economic conflicts of interest. Studies have demonstrated, for instance, that auditors are more lax when they monitor their own paying clients (Duflo, Greenstone, Pande, and Ryan, 2013; Jiang, Stanford, and Xie, 2012; Kinney, Jr., Palmrose, and Scholz, 2004), when they face more competition (Bennett *et al.*, 2013), and when they operate in corrupt institutional environments in which they are more likely to receive side payments from audited firms (Montiel *et al.*, 2012).

While these studies identify important criteria affecting the quality of information generated through audits, we seek to move beyond economic incentives to investigate the influence of social institutions and relationships. Such influence is well established in the literature on public regulatory enforcement (Brehm and Gates, 1997; Keiser and Soss, 1998; Lipsky, 1980/2010; May and Winter, 2000; Maynard-Moody and Musheno, 2003; Piore, 2005; Piore and Schrank, 2008; Sandfort, 2000; Schrank, 2009), but has not been explored in the context of private monitoring. This is a significant gap in the literature, particularly as private third-party auditors become increasingly central both to effectuating global firms' risk management strategies and to enforcing domestic and transnational regulatory regimes. The assumption in the literature has been that the profit motive is the dominant, if not exclusive, influence driving private monitors (Bazerman, Morgan, and Loewenstein, 1997; Cantor and Packer, 1994; Dallas, 2011; Moore *et al.*, 2006; Manns, 2013; Oh, 2004; Partnoy, 2006). While

we recognize that private auditors are subject to very different incentive structures than government inspectors, we resist the premise that economic incentives will entirely crowd out social influences, and we draw on the literature on regulatory compliance to theorize how social as well as economic factors will influence private auditors' decisions. These insights are crucial to understanding how firms can manage their monitors and maximize the validity of the strategic information they receive from their auditors about their suppliers.

HYPOTHESES

Ongoing auditor-supplier relationships

Many have theorized that auditors who repeatedly visit an audited entity are likely to be subject to cognitive biases and social pressures that will influence which violations they detect and cite. Bounded rationality limits the number of issues an auditor can pursue during any given audit (Jones, 2001; Simon, 1947). As Chugh and Bazerman (2007: 3) have argued, "bounded awareness" causes individuals 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).

These cognitive biases may be reinforced by social pressures. Returning auditors may develop "cozy relationships" (Moore *et al.*, 2006: 24) with an audited firm's management that leads them to identify with its positions. Familiarity between auditors and management may also foster corruption, emboldening managers to pressure auditors to report good results.

Empirical research has documented how managers' awareness is bounded by their past experiences. Specifically, "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, Roos, and Slocum, 1994: 58) and their “perceptual and cognitive limitations” have been shown to lead to errors (Huber and Power, 1985: 172). Henderson and Clark (1990) similarly demonstrate that professional engineers who advise management tend to approach new problems through the lens of their experience solving previous problems, restricting their ability to identify innovative solutions. Empirical research on government inspectors has found that ongoing relationships between inspectors and inspected entities encourage a “benefit of the doubt” style of enforcement rather than an arms-length “policing” style (Bardach and Kagan, 1982/2002) and result in less severe penalties (Muehlenbachs, Staubli, and Cohen, 2013).

Our interviews with social auditors revealed similar concerns that an auditor who keeps auditing the same facility may “go native,” becoming an extension of the supplier’s management. In addition, we have no reason to believe that private auditors will not be subject to the same bounded awareness constraints as other monitors. Returning to the same supplier, they are likely to focus on the domains they highlighted previously, whereas a completely new audit team would examine a supplier with a fresh set of eyes and focus on a different set of issues, likely uncovering new violations.

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

Auditor tenure

Scholars and activists have suggested that more experienced supply chain auditors are more effective (Esbenshade, 2004; Locke, Qin, and Brause, 2007). It is not clear, however, how experience affects the number of violations cited in a given audit. On the one hand, experience 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 supplier uses 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, although such marginal gains later diminish.

On the other hand, it is not clear that more experienced auditors will *cite* more violations. 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” posed by particular violations, so they tend to go by the book and cite everything (Bardach and Kagan, 1982/2002: 129). Experienced inspectors, by contrast, may decline to cite violations lacking the requisite level of risk and culpability (Bardach and Kagan, 1982/2002; Hawkins, 1984). We therefore expect that 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 about which violations to cite.

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

Professionalization

Education and training should promote both detection and citation of violations (Chen, Chang, and Lee, 2008). More professionalized auditors may feel more obligated to cite what they find. Sociologists have long theorized that professionalization—specialized education and training in a field’s skills and values—is a key constraint on individual discretion in both corporate and government bureaucracies (Abbott, 1988; Freidson, 1994; Larson, 1977; Scott, 1966). 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 preserving and transmitting shared values.

More recently, Lipsky (1980/2010: 201) argued 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 influence of economic incentives on 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 (Gordon and Simon, 1992; Lipsky, 1980/2010; Thompson, 1967), 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). We therefore expect that teams whose auditors are more professionalized will record more violations.

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

Gender

Research has suggested that, even when constrained by bureaucratic rules and roles, men and women may perform their work “somewhat differently” (Eagly and Johannesen-Schmidt, 2001: 783). Several gender-based behavioral distinctions documented in the literature can

influence whether social auditors discover—and then, whether they cite—violations. Research has shown that women are more persistent at pursuing assigned tasks (Jacob, 2002; Spence and Buckner, 2000; Stonewater, Eveslage, and Dingerson, 1990), suggesting more diligent enforcement of regulations. Research has also found that women have perceptual and integrative processing advantages that may enhance their ability to detect violations. For example, women have been found to be more skilled at interpreting the emotional content of others' expressions (Campanella *et al.*, 2004; Killgore and Cupp, 2002; Thayer and Johnsen, 2000) 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, 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 better equip them to perceive violations in a complex factory setting and to elicit information about violations from employees.

Moreover, research suggests that women are more likely to cite the violations they perceive. Women in bureaucratic organizations are more likely than men to be strict rule-followers (Oberfield, 2010; Portillo, 2012; Portillo and DeHart-Davis, 2009). 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 (Green and Melnick, 1950; Kanter, 1977; Thompson, 1977). We are not aware of any research on women's status in supply chain auditing, but research on financial auditors and audit firms suggests that, even as many

women have entered that profession, audit firms have maintained masculine organizational cultures that tend to devalue women's contributions (Haynes, 2012; Jonnergård, Stafsudd, and Elg, 2010; Mueller, Carter, and Ross-Smith, 2011). Our interviews with social auditors suggest that, especially in societies with more rigid gender hierarchies, supplier managers view male auditors as more authoritative. Empirical studies of government workers have found that women do indeed "go by the book" (Green and Melnick, 1950; Portillo, 2012: 90; Portillo and DeHart-Davis, 2009) more strictly than their male colleagues do. All this evidence suggests that gender will significantly influence whether supply chain auditors 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 gender diversity will enhance a supply chain auditing team's 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 violation, enabling mixed-gender teams to find more. This should be particularly valuable in eliciting information from a diverse set of supply chain 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, while, as one interviewee put it, male

supplier managers, “for cultural reasons, may find it difficult...to open up to women.”

Second, research has shown how the interpersonal dynamics of gender diversity can improve team performance. For instance, studies find that people on socially diverse teams tend to prepare more thoroughly and to think through a broader range of issues (Loyd *et al.*, 2013). Fenwick and Neal report the superior performance of gender-diverse teams at management-simulation exercises, crediting it 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 see they are being outperformed by women. Studies have shown that lower-performing team members often compare themselves to better performers, which tends to raise their “usual performance levels in order to match or beat the stronger performers” (Collins, 2000; Lount, Jr. and Phillips, 2007; Weber, Wittchen, and Hertel, 2009: 732). 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 their female teammates’ citation rates.

Although some conflicting evidence suggests that gender diversity can sometimes undermine team performance, 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 interact a lot with the individuals they audit, we expect that gender diversity will improve team performance.

Hypothesis 5 (H5): Audits conducted by gender-diverse 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 for thousands of code-of-conduct audits

conducted in 66 countries between 2004 and 2009 by one of the world's largest social auditing companies.¹ During that period, the company, which already had more than a decade's experience, employed several hundred people in many countries; they spoke over 30 languages.

The dataset contains audit results for and information about each audited supplier, including its country and a unique identifier; characteristics and unique identifiers for the auditors on each audit; and the country of the multinational firm on whose behalf each audit was conducted and a unique identifier for that firm. The auditor preserved the anonymity of the factories, auditors, and multinational firms by not revealing their names. Our estimations are based on the 16,795 audits of 5,819 factories (in 66 countries) for which we had data on all the measures described below and which had been audited at least twice during the sample period (a technical requirement owing to our models being estimated with supplier-level fixed effects, described below). The country and industry composition of our sample is reported in Tables 1 and 2; the most common industries are garments, accessories, electronics, and toys.

[Insert Tables 1 and 2 about here]

In nearly all cases, multinational firms determined (a) which suppliers would be audited and (b) whether the multinational firm or the supplier (or its agents) would pay. Our interviews indicated that this decision was not driven by the supplier'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.)

¹ The company required anonymity as a condition of sharing its data with us.

Dependent variable

We measure the extent to which factories adhere to codes of conduct as the *number of violations* in each audit,² obtained from the social auditing firm's database. We include only the types of violation that, according to the social auditing firm (hereafter referred to as "the auditing firm"), apply in all industries and 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.³ During an audit, the auditors code a common set of dichotomous indicators (violation or no violation) in each category.⁴

Independent variables

To identify the potential for auditors' career concerns to influence their behavior, we coded *previous auditor* as 1 when at least one member of the focal audit team had participated in one of the supplier's previous audits during the sample period and 0 otherwise.

We measure an auditor's experience as his or her years of service at the auditing firm based on data from the 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 the maximum tenure yielded nearly identical results.)

We measure the professionalism of the audit team in two ways. Because one important

² Studies of compliance with government health and safety regulations have long used violation counts recorded by inspectors as a measure of compliance variation (Braithwaite and Makkai, 1991; Gray and Scholz, 1993; Gray and Shadbegian, 2005) and more recent studies of supplier compliance with private labor standards have used violation data recorded by private auditors (Ang *et al.*, 2012; Oka, 2010).

³ 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.

⁴ The occupational health and safety category, for example, 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.

source of professionalization is “standardized formal training in universities” (Lipsky, 1980/2010: 201), we code *graduate education* as 1 when at least one member of the audit team had a graduate degree and 0 otherwise. We focused on graduate education because nearly all auditors in our dataset had a bachelor’s degree.⁵ We also created *auditing skills training* as the highest number of the auditing firm’s training courses that any audit team member had completed. These courses teach 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 yielded nearly identical results.)

We measure gender composition with three dummy variables—*all-female audit team*, *all-male audit team*, and *mixed-gender audit team*.⁶

Control variables

Training can influence the stringency of government regulators (Macher, Mayo, and Nickerson, 2011). We therefore control for two types of training that might influence either an audit team’s ability to detect and report violations or the nature of the violations it detects and reports. Using the auditing firm’s database, we calculated the proportion of each team that had undergone *certification training*—training on the standards and protocols of a particular certification regime, such as SA8000—because the firm indicated that such training influences the scope of the audit and the types of violation auditors look for. We also calculated the proportion of each team that had undergone *brand training*—training provided by the multinational firm on its corporate responsibility program and procedures—to account for

⁵ 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.

⁶ We use these dummies rather than a continuous measure such as proportion female because the database indicated that 97 percent of the audit teams in our sample were all-female, all-male, or evenly divided. Thus, the three dummies represent the distribution of our data.

possible variations in stringency requested by different brands.

We control for auditors' age to ensure that the effects of auditor tenure can be attributed to job experience rather than to 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 auditing firm provided only five-year age-range categories (for example, 20–24 years old) for each auditor to keep precise ages confidential, we created a proxy for the team's average age. We calculated the midpoint for each category and then created *average age* as the average of the oldest and youngest age-range categories on a team. (Using the oldest team member's age rather than the average yielded nearly identical results.)

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 Worldwide Responsible Accredited Production (WRAP)—because such protocols might limit an auditor's discretion.

We also control for whether an audit is unannounced or pre-announced, using a dummy variable, *unannounced audit*, coded 1 for an unannounced audit and 0 for a pre-announced audit. The latter provides several weeks of notice, giving the supplier time to try to remedy problems, which could result in fewer violations to find.

Because research has indicated that financial conflicts of interest created by audit fees undermine auditors' and inspectors' stringency in policing corporate misconduct (Bazerman, Morgan, and Loewenstein, 1997; Cantor and Packer, 1994; Dallas, 2011; Duflo *et al.*, 2013; Estlund, 2012; Manns, 2013; Moore *et al.*, 2006; Oh, 2004; Partnoy, 2006; Pierce and Toffel, 2013), we control for which entity paid for each audit. Using the auditing firm's database, we

created two dichotomous variables. *Audit paid for by supplier or agent* identifies audits with the potential for financial conflict of interest. It is coded 1 for audits paid for by the audited supplier or by agents, vendors, or licensees and coded 0 for audits paid for by the multinational firm.⁷ *Audit paid for by the multinational firm* is coded in the opposite manner.

We include dummy variables to control for the *number of auditors* on each audit (two through five, with one as the omitted category). Audit team size is a direct function of supplier size and complexity in our setting, but others have shown that larger teams of government inspectors can lead to more stringent monitoring (Muehlenbachs, Staubli, and Cohen, 2013).

We created a dummy variable *re-audit* to distinguish routine audits from re-audits, which tend to focus on domains where violations were previously identified and which therefore tend to yield fewer violations than routine audits. We also include a series of dummies to indicate a supplier's audit sequence—its *second audit*, *third audit*, and so on through *sixth-or-higher audit* (because only five percent of the audits in our sample were a supplier's seventh or higher audit), with a supplier's first audit as the omitted category—to control for the possibility that successive audits yield fewer violations as factories address the issues exposed. (Using an *audit sequence* counter variable and its square rather than the dummies yielded nearly identical results.)

To capture domestic institutional factors that could influence a supplier's compliance with codes of conduct (Toffel, Short, and Ouellet, 2014), we control for several country-level governmental, economic, and civil-society attributes. We measure the average economic development level of a supplier's country in the year the audit was conducted as its annual *per-*

⁷ We combined these categories of payer 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, 2011) 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 for by factories as distinct from audits paid for by agents, vendors, or licensees.

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 use the log. To measure the extent to which the government of the supplier’s country fosters a regulatory environment promoting economic development, we use the annual *regulatory quality* metric corresponding to the year the audit was conducted. This metric is calculated by the World Bank’s Worldwide Governance Indicators project (obtained from <http://data.worldbank.org/data-catalog/worldwide-governance-indicators>) 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).⁸

We measure the extent of *press freedom* in the supplier’s country the year the audit took place 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 create annual *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 3 and 4.

[Insert Tables 3 and 4 about here]

EMPIRICAL MODEL AND RESULTS

Our empirical model includes all independent and control variables described above and three sets of fixed effects. Specifically, we estimate the following model:

⁸ Controlling instead for supplier countries’ annual Corruption Perceptions Index scores from Transparency International, which are highly correlated with the World Bank’s *regulatory quality* metric ($\rho=0.96$), yields nearly identical results.

$$Y_{icdj} = F(\beta_1 X_{id} + \beta_2 \phi_{id} + \beta_3 \lambda_{cd} + \beta_4 \alpha_i + \beta_5 \delta_t + \beta_6 \gamma_j + \varepsilon_{icdj}),$$

where Y_{icdj} refers to the *number of violations* recorded in the audit of supplier i in country c that was conducted on date d on behalf of a multinational firm in country j . $F(\cdot)$ refers to the Poisson function. X_{id} refers to our hypothesized variables (*previous auditor, maximum tenure, average tenure, graduate education, auditing skills training, all-male audit team, all-female audit team, mixed-gender audit team*). ϕ_{id} refers to the audit-level control variables described above (*certification training, brand training, average age, maximum age, third-party protocol, unannounced audit, audit paid for by supplier or agent, audit paid for by multinational firm, re-audit*, fixed effects denoting *number of auditors* [two through five-or-more], and fixed effects indicating the supplier's *audit sequence* [second through sixth-or-more]). λ_{cd} refers to the annual supplier-country control variables described above (*per-capita GDP (log), regulatory quality, press freedom*). We include fixed effects for each supplier (α_i) to control for time-invariant characteristics that might affect that factory's violation rate, such as size, age, industry, and national institutional context.⁹ A series of dummies for the year in which the audit was conducted (δ_t) controls for overall temporal trends. We also include fixed effects for the headquarters country of the multinational firm on whose behalf each audit was conducted (γ_j). This controls for the possibility that consumers and activist groups in different multinational firm countries vary in their concern for and attentiveness to supply chain conditions, which might in turn affect how much pressure firms exert on their supply chain auditors to audit stringently. These fixed effects also control for all other time-invariant differences between the multinational firm headquarters countries' institutional contexts.

⁹ Because supplier-level fixed effects in our model absorb the time-invariant portion of supplier-country-level variables, *per-capita GDP*, *regulatory quality*, and *press freedom* effectively control for within-country temporal variation in their effect on supplier violation rates.

Our identification strategy is based on the fact that the process of assigning auditors to audit teams is unrelated to our independent variables and thus is not a source of endogeneity that should bias their coefficients. Specifically, our interviews with the social auditing firm indicated that assignments were based on (1) language skills to communicate with management and workers, (2) availability, and (3) the need for at least one team member to qualify as a lead auditor. Potential concerns that endogenous audit assignment—and, in particular, differences between lead auditors and other auditors—might bias our results led us to conduct several supplemental analyses. As described in detail in the Appendix, our comparison of lead and non-lead auditors’ demographic characteristics (such as languages spoken, age, experience, and training) and estimation of instrumental variable regressions yielded no evidence indicating endogeneity bias.

Results

We estimate the model using Poisson regression with robust standard errors and report our results in Column 1 of Table 5. Negative-binomial regression with conditional fixed effects yields nearly identical results, indicating that our results are not sensitive to estimation technique. Variance inflation factors (VIF) were below 2 for all hypothesized variables and below 5 for all variables except a few buyer-country dummies, which yields no evidence that multicollinearity is a serious concern. The explanatory power of the model is indicated by a McFadden’s R-squared value of 0.39 and a McFadden’s adjusted R-squared value of 0.30.

[Insert Table 5 about here]

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 a team member had participated in a prior audit at the same supplier, which supports Hypothesis 1. The average marginal effect

(AME) of -0.3 indicates that, compared to the sample average of 6.5 violations, an audit by a team with a previous auditor 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 cited increases as tenure increases but at a diminishing rate, which supports Hypothesis 2. This relationship is illustrated in Figure 1, which graphs average predicted violations at varying levels of the audit team's maximum experience.

[Insert Figure 1 about here]

Our results are mixed with respect to professionalization. The coefficient on *graduate education* is positive as predicted but not significant, yielding 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 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 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, but only when professionalization is measured by specific training rather than by broader education.

Team gender composition is also significantly associated with the number of violations reported. Audits by all-female teams yield 6 percent more violations than those by all-male teams (the baseline) ($\beta = 0.05$; $p < 0.01$; IRR = 1.05), which supports Hypothesis 4. The average

marginal effect indicates that audits by all-female teams yield 0.35 more violations than those by all-male teams (the baseline category).

Mixed-gender teams 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) and slightly more violations than all-female teams (Δ AME = 0.1), but the latter difference is not statistically distinguishable ($\chi^2 = 0.55$; $p = 0.46$). These results partially support Hypothesis 5.

The coefficients on *second inspection* through *sixth or more inspection* are negative and statistically significant. Wald tests comparing these coefficients indicate that, on average, each successive audit yields significantly fewer violations. Average marginal effects indicate that, on average, a supplier'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 nearly 1.3 fewer 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 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 assigning more auditors to larger factories, which are likely to generate more violations, we find that audits with more auditors yield significantly more violations. We find no evidence that the number of violations varied with the team's certification training, brand training, or average age or with a third-party protocol. Our point estimate indicates that unannounced audits yielded slightly more violations than announced audits at a given supplier (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 the multinational firm, the baseline category ($\beta = -0.08$; $p < 0.01$; IRR = 0.92). The average marginal effect indicates that, on average, audits yield 0.6 fewer violations when the supplier or agent pays than when the multinational firm pays, a drop from 6.5 to 5.9.

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

Robustness tests

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 member rather than the average of the oldest and youngest members' ages) and an alternative approach to controlling for the supplier's audit sequence (a counter and its square instead of dummies), the results were nearly identical to our primary results. As described in the Appendix, instrumenting for the audit team's maximum *audit skills training* and *maximum tenure* using average values of these characteristics among all auditors based in the auditing firm's field office that staffed each establishment's audits yielded statistically indistinguishable results (Hausman test $\chi^2 = 52.09$; $p = 0.16$).

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 5 reports estimates after excluding the 210 audits performed for multinational firms whose audit teams were always all-female, in case that pattern reflected a multinational firm's policy that might bias our primary results. Column 3 reports estimates based on the 10,648 audits conducted by teams of at least

two members to ensure that our results were not driven by single-auditor audits. 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, in these cases, the factories themselves might have chosen the protocol and auditor. Column 5 reports results for the subsample of 9,266 audits that excludes each supplier's first audit in our sample; some of those might have been pre-assessments of factories that multinational firms had not yet engaged and our hypothesized relationships might operate differently in such circumstances. 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 theorized and tested several social factors that shape how supply chain auditors identify and report violations of supplier codes of conduct. We find that auditors' decisions are shaped not only by the financial conflicts of interest that have been the focus of research to date, but also by social factors, including the auditors' experience, professional training, and gender; the gender diversity of their teams; and their repeated interactions with those whom they audit. These findings contribute to several literatures and suggest strategies for designing private monitoring regimes to provide companies with more reliable strategic information.

Contributions to the strategic management literature

Our study significantly extends the TCE literature's conceptions of monitoring to address key strategic concerns associated with the global outsourcing of production. First, we highlight the particular importance of supply chain monitoring to mitigate transaction costs that can arise

from negative reputational spillovers. To our knowledge, the TCE literature has not explored this very substantial potential cost of outsourcing to suppliers with poor social or environmental practices. In doing so, we bridge the TCE literature with strategic management perspectives on reputation (Roberts and Dowling, 2002). Second, we add an important empirical dimension to the largely theoretical TCE literature on monitoring. Our empirical documentation of heterogeneity in the quality of audit information suggests that any analyses of transaction costs mitigated by monitoring must consider not only the existence but the efficacy of different monitoring structures. Third, we extend the literature's focus on the *ex ante* design of monitoring structures to highlight the importance of ongoing, *ex post* monitoring in contractual relationships. Finally, we identify important second-order monitoring problems that arise when firms outsource their monitoring functions to third-party auditors and our findings offer strategies firms can use to “monitor their monitors” and increase the reliability of the strategic information generated in supply chain audits.

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, the practices of the auditors themselves have largely remained a black box. We illuminate 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 when those being audited are paying for the audits. Our analysis indicates that while economic incentives do play a role, private-sector auditor behavior is also significantly influenced by social institutions, identities, and relationships. Our finer-grained picture suggests that audit designers should moderate potential bias and increase audit reliability by considering

the auditors' characteristics and relationships that we found significantly influence their decisions.

Our findings should likewise inform the broader literature on private gatekeepers such as accountants and credit rating agencies, subjects of much interest 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, the gatekeeper literature, like the auditing literature, has focused almost exclusively on the influence of economic conflicts of interest (Bazerman, Morgan, and Loewenstein, 1997; Cantor and Packer, 1994; Dallas, 2011; Goldberg, 1988; Hill, 2004; Manns, 2013; Moore *et al.*, 2006; Oh, 2004; Partnoy, 2006; Schwarcz, 2002). Our study suggests the need to look at a broader range of factors in order to structure more 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 needed dialogue between the 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, while research on public-sector monitors has largely ignored front-line policy 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. For example, while several studies of government monitors—including patent examiners (Lemley and Sampat, 2012) and U.S. Food and Drug Administration inspectors

of pharmaceutical plants (Macher, Mayo, and Nickerson, 2011)—have found an inverse relationship between experience and inspection stringency, we found that private monitors’ stringency increases as experience increases, although at a declining rate.

Contributions to research on transnational business regulation

Supply chain auditing has become an important component of international regulatory schemes that seek to address the social and environmental risks of global business activities (Braithwaite and Drahos, 2000; Scott, 2012). Private labeling regimes such as the Forest Stewardship Council, the Marine Stewardship Council, and Fair Trade rely on private third-party auditors. International intergovernmental institutions such as the United Nations have encouraged supply chain auditing by requesting that MNCs conduct “due diligence” to ensure their suppliers’ compliance with international human rights norms (Kamatali, 2012; Ruggie, 2008; Shamir, 2005). Many national regulators have followed suit, requiring MNCs to conduct due diligence and disclose supply chain practices (Zandvliet, 2011). The efficacy and legitimacy of private voluntary transnational business regulation largely depends on the credibility of monitoring; our study responds to calls for more empirical research on the key actors (Büthe, 2010). While our findings of auditor heterogeneity support those who question auditor independence and objectivity (Boiral and Gendron, 2011; Power, 1997), our identification of several systematic determinants of that heterogeneity suggests how companies and policymakers can improve audit validity. In addition, as governments begin to mandate certain ESG measures and disclosures that were once entirely private and voluntary, our findings suggests how firms can mitigate legal costs and risks associated with their practices in these domains.

Implications for managers

Our study provides tools to help firms manage their monitors to produce complete and

accurate information about suppliers that will enable them to make better strategic decisions about which suppliers to work with. Our finding that auditors tend to cite fewer violations at factories where they have ongoing relationships empirically supports managers who advocate auditor rotation to deter capture (Moore *et al.*, 2006; U.S. Public Company Accounting Oversight Board, 2011). That auditors returning to the same supplier might consistently deemphasize some areas (and thus inadvertently overlook violations) should also encourage multinational firms to ensure that the auditors they hire engage in routine auditor rotation.¹⁰ Our findings also highlight the importance of companies ensuring that the teams that audit their suppliers are well trained. While auditors with higher educational credentials did not find significantly more supplier violations than less-educated peers, those with more audit-specific training did find significantly more supplier violations than those with less audit-specific training. Finally, our findings suggest that firms should ensure that the audit teams assessing their suppliers have sufficient experience and gender diversity.

Limitations and 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 employees of the firm that provided our data and employees of competing firms, yielded no reason to suspect that endogeneity concerns are driving our results. These discussions indicated that audit team assignments were driven largely by language skills, availability, and the need for each team to have a qualified lead auditor. Our discussions also indicated that multinational firms determine which factories are

¹⁰ A few auditing schemes have explicitly stipulated term limits for auditing companies; for example, California's greenhouse gas regulation requires regulated entities to change verification companies every six years.

audited, which obviates the risk of a selection effect whereby better-than-average or worse-than-average factories might choose to be audited or to pay for their own audits, as does happen in some voluntary environmental programs (King and Toffel, 2009). Even so, we cannot rule out the possibility that omitted variables are correlated with our independent variables and violation rates; we therefore encourage future randomized field experiments (e.g., Hainmueller, Hiscox, and Sequeira, 2011).

Our findings relating to gender and gender diversity may be influenced by the gender composition of the supplier's workforce. Although we do not have such demographic data, available meta-data and our own interviews with social auditors suggest that women dominate the workforce in export-intensive industries such as garments, textiles, and electronics, which account for most of our sample (Dejardin and Owens, 2009; Jenkins, Esquivel, and Larrían, 2001; Kuncoro, 2011). Future research could explore how auditors' decisions are influenced by the interaction of the gender compositions of the audit team and the audited organization.

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. Field experiments might shed light on which types of technical and managerial training most improve auditors' objectivity. More broadly, it will be important to investigate whether our findings are generalizable to different types of private gatekeepers, such as financial auditors, credit ratings agencies, and attorneys. Do they respond similarly to economic incentives, professional obligations, and social pressures? In addition, direct comparisons of the implementation practices of private-sector monitors such as social auditors and public-sector monitors such as government inspectors could reveal opportunities to

enhance the efficiency and effectiveness of both. Finally, from the TCE perspective, it is important to investigate whether social monitoring actually mitigates reputational transaction costs for outsourcing firms, either by improving the social practices of suppliers or by providing firms with timely actionable information about reputational risks in their supply chains.

CONCLUSION

Although private supply chain auditors are increasingly important to strategic corporate outsourcing decisions and to public and private transnational business regulation, 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 strategic management, private supply chain monitoring, and regulatory compliance mechanisms and highlights opportunities to improve the design and implementation of monitoring outsourced production.

REFERENCES

- Abbott, A. 1988. *The System of Professions: An Essay on the Division of Expert Labor*. Chicago: University of Chicago Press.
- 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. 2014. Learning by supplying. *Strategic Management Journal* 35(2): 204–223.
- Ang, D., D. Brown, R. Dehejia, and R. Robertson. 2012. Public disclosure, reputation sensitivity, and labor law compliance: Evidence from Better Factories Cambodia. *Review of Development Economics* 16(4): 594–607.
- Argyres, N. S., J. Bercovitz, and K. J. Mayer. 2007. Complementarity and evolution of contractual provisions: An empirical study of IT services contracts. *Organization Science* 18(1): 3–19.
- 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.
- Bardach, E., and R. A. Kagan. 1982/2002. *Going by the Book: The Problem of Regulatory Unreasonableness*. New Brunswick, NJ: Transaction.
- Barthélemy, J., and B. V. Quélin. 2006. Complexity of outsourcing contracts and *ex post* transaction costs: An empirical investigation. *Journal of Management Studies* 43(8): 1775–1797.
- 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.
- Bennett, V. M., L. Pierce, J. A. Snyder, and M. W. Toffel. 2013. Customer-driven misconduct: How competition corrupts business practices. *Management Science* 59(8): 1725–1742.
- Black, J. 1997. *Rules and Regulators*. Oxford, UK: Clarendon.
- Blair, M., C. Williams, and L. Lin. 2008. The new role for assurance services in global commerce. *Journal of Corporation Law* 33(2): 325–360.
- Boiral, O. 2007. Corporate greening through ISO 14001: A rational myth? *Organization Science* 18(1): 127–146.
- Boiral, O., and Y. Gendron. 2011. Sustainable development and certification practices: Lessons learned and prospects. *Business Strategy and the Environment* 20(5): 331–347.
- Bowen, F. 2014. *After Greenwashing: Symbolic Corporate Environmentalism and Society*. Cambridge, UK: Cambridge University Press.
- Braithwaite, J., and P. Drahos. 2000. *Global Business Regulation*. Cambridge, UK: Cambridge University Press.
- Braithwaite, J., and T. Makkai. 1991. Testing an expected utility model of corporate deterrence. *Law & Society Review* 25(1): 7–40.
- 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(Summer-Fall): 1–26.
- Card, D. 1995. Using geographic variation in college proximity to estimate the return to schooling. In L. N. Christofides, E. K. Grant, and R. Swidinsky (eds.), *Aspects of Labor Market Behaviour: Essays in Honour of John Vanderkamp*. Toronto: University of Toronto Press.
- 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(3): 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(6): 863–883.
- Chugh, D., and M. H. Bazerman. 2007. Bounded awareness: What you fail to see can hurt you. *Mind and Society* 6(1): 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.
- Crilly D., and P. Sloan, 2012. Enterprise logic: Explaining corporate attention to stakeholders from the ‘inside-out.’ *Strategic Management Journal* 33(10) 1174–1193.
- 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(3): 343–358.
- Delmas, M. 2002. The diffusion of environmental management standards in Europe and in the United States: An institutional perspective. *Policy Sciences* 35(1): 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* 36(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(3): 189–196.
- Dogui, K., O. Boiral, and Y. Gendron. 2013. ISO auditing and the construction of trust in auditor independence. *Accounting, Auditing & Accountability Journal* 26(8): 1279–1305.
- Doig, S. J., R. C. Ritter, K. Speckhals, and D. Woolson. 2001. Has outsourcing gone too far? *McKinsey Quarterly* 4: 25–37.
- Duflo, E., M. Greenstone, R. Pande, and N. Ryan. 2013. Truth-telling by third-party auditors and the response of polluting firms: Experimental evidence from India. *Quarterly Journal of Economics* 128(4): 1499-1545.
- 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.
- Eisenhardt, K. M. 1985. Control: Organizational and economic approaches. *Management Science* 31(2): 134–149.
- Ellis, P. D. 2011. International trade intermediaries. In J. N. Sheth and N. K. Malhotra (eds.), *Wiley International Encyclopedia of Marketing, Part 6. International Marketing*. West Sussex, UK: Wiley-Blackwell.
- Esbenshade, J. 2004. *Monitoring Sweatshops: Workers, Consumers and the Global Apparel Industry*. Philadelphia: 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(1): 115–154.
- Feinstein, J. S. 1990. Detection controlled estimation. *Journal of Law and Economics* 33(1): 233–276.
- Fenwick, G. D., and D. J. Neal. 2001. Effect of gender composition on group performance. *Gender, Work & Organization* 8(2): 205–224.
- Forman, C., A. Goldfarb, and S. Greenstein. 2012. The Internet and local wages: A puzzle. *American Economic Review* 102(1): 556–575.
- Freidson, E. 1994. *Professionalism Reborn: Theory, Prophecy, and Policy*. Chicago: University of Chicago Press.
- Fung, A., D. O'Rourke, and C. Sabel, 2001. *Can We Put an End to Sweatshops?* Boston: Beacon Press
- Garicano, L., and T. N. Hubbard. 2009. Specialization, firms, and markets: The division of labor within and between law firms. *Journal of Law, Economics, and Organization* 25(2): 339–371.
- 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. Goma. 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.
- Gray, W. B., and J. T. Scholz. 1993. Does regulatory enforcement work? A panel analysis of OSHA enforcement. *Law & Society Review* 27(1): 177–214.
- Gray, W. B., and R. Shadbegian. 2005. When and why do plants comply? Paper mills in the 1980s. *Law & Policy* 27(2): 238–261.
- 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.
- Greenhouse, S. 2013a. \$40 million in aid set for Bangladesh garment workers. *New York Times*, December 24: B1.
- Greenhouse, S. 2013b. Major retailers join Bangladesh safety plan. *New York Times*, May 13: B1.
- Guler, I., M. F. Guillén, 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(2): 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, Cambridge, MA.
- 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.

- Heide, J. B., A. Kumar, and K. H. Wathne. 2014. Concurrent sourcing, governance mechanisms, and performance outcomes in industrial value chains. *Strategic Management Journal* 35(8): 1164–1185.
- Heide, J. B., K. H. Wathne, and A. I. Rokkan. 2007. Inter-firm monitoring, social contracts, and relationship outcomes. *Journal of Marketing Research* 44(3): 425–433.
- 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(1): 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(2): 171–180.
- Hult, G. T. M., D. J. Ketchen, and M. Arrfelt. 2004. Information processing, knowledge development, and strategic supply chain performance. *Academy of Management Journal* 45(3): 577–586.
- Hult, G. T. M., D. J. Ketchen, and M. Arrfelt. 2007. Strategic supply chain management—Improving performance through a culture of competitiveness and knowledge development. *Strategic Management Journal* 28(10): 1035–1052.
- Ioannou, I. 2014. Redefining strategy in the age of sustainability and social responsibility. Available at SSRN: <http://ssrn.com/abstract=2429870>.
- 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. Larrián. 2001. Export processing zones in Central America. In F. Larraín (ed.), *Economic Development in Central America, Vol. I: Growth and Internationalization*. Cambridge, MA: Harvard University.
- Jiang, J. (X.), M. H. Stanford, and Y. Xie. 2012. Does it matter who pays for bond ratings? Historical evidence. *Journal of Finance Economics* 105(3): 607–621.
- Jones, B. D. 2001. *Politics and the Architecture of Choice: Bounded Rationality and Governance*. Chicago: University of Chicago Press.
- Jonnergård, K., A. Stafssudd, 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.
- Kalnins, A., and K. J. Mayer. 2004. Relationships and hybrid contracts: An analysis of contract choice in information technology. *Journal of Law, Economics, and Organization* 20(1): 207–229.
- 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.

- Kayser, S. A., J. W. Maxwell, and M. W. Toffel. 2014. Supply chain screening without certification: The critical role of stakeholder pressure. Working Paper 15-009, Harvard Business School, Boston.
- 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., 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.
- 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, B., and S. A. Soule. 2007. Social movements as extra-institutional entrepreneurs: The effect of protest on stock price returns. *Administrative Science Quarterly* 52(3): 413–442
- Kinney, W. R., Jr., Z. Palmrose, and S. Scholz. 2004. Auditor independence, non-audit services, and restatements: Was the U.S. government right? *Journal of Accounting Research* 42(3): 561–588.
- 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., F. Qin, and A. Brause. 2007. Does monitoring improve labor standards? Lessons from Nike. *Industrial and Labor Relations Review* 61(1): 3–31.
- 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, Cambridge, MA.
- Locke, R. M., and M. Romis. 2007. Improving work conditions in a global supply chain. *MIT Sloan Management Review* 48(2): 54–62.
- 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(2): 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(3): 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.
- Mayer, K. J., and R. M. Salomon. 2006. Capabilities, contractual hazards, and governance: Integrating resource-based and transaction cost perspectives. *Academy of Management Journal* 49(5): 942–959.
- 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.
- Meidinger, 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(9): 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.
- Muehlenbachs, L., S. Staubli, and M. A. Cohen. 2013. The effect of inspector group size and familiarity on enforcement and deterrence. Discussion paper 13–36, Resources for the Future, Washington, DC.
- 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.
- 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.
- O'Callaghan, T. 2007. Disciplining the multinational enterprise: The regulatory power of reputation risk. *Global Society* 21(1): 95–117.
- Oh, P. B. 2004. Gatekeeping. *Journal of Corporation Law* 29(4): 735–800.
- Oka, C. 2010. Accounting for the gaps in labour standard compliance: The role of reputation-conscious buyers in the Cambodian garment industry. *European Journal of Development Research* 22(1): 59–78.
- 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.
- Park, S. K. 2014. Targeted social transparency as global corporate strategy. *Northwestern Journal of International Law & Business* 34.
- 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(3): 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(2): 339–347.
- Potoski, M., and A. Prakash. 2004. Regulatory convergence in nongovernmental regimes: Cross-national adoption of ISO 14001 certification. *Journal of Politics* 66(3): 885–905.
- Potoski, M., and A. Prakash. 2005a. Covenants with weak swords: ISO 14001 and facilities' environmental performance. *Journal of Policy Analysis and Management* 24(4): 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, M. 1997. *The Audit Society: Rituals of Verification*. Oxford, UK: 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.
- Puranam, P., R. Gulati, and S. Bhattacharya. 2013. How much to make and how much to buy? An analysis of optimal plural sourcing strategies. *Strategic Management Journal* 34(10): 1145–1161.
- Reitzig, M., and S. Wagner. 2010. The hidden costs of outsourcing: Evidence from patent data. *Strategic Management Journal* 31(11): 1183–1201.
- Roberts, P. W., and G. R. Dowling. 2002. Corporate reputation and sustained superior financial performance. *Strategic Management Journal* 23(12): 1077–1093.
- 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. <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 and C. A. Rodríguez-Garavito (eds.), *Law and Globalization from Below*: 92–117. Cambridge, UK: 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(1): 44–62.
- Stonewater, B. B., S. A. Eveslage, and M. R. Dingerson. 1990. Gender differences in career helping relationships. *Career Development Quarterly* 39(1): 72–85.
- 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(1): 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., and J. L. Short. 2011. Coming clean and cleaning up: Is voluntary disclosure a signal of effective self-policing? *Journal of Law and Economics* 54(3): 609–649.
- Toffel, M. W., J. L. Short, and M. Ouellet. 2014. Codes in context: How states, markets, and civil society shape adherence to global labor standards. Working Paper 13–045, Harvard Business School, Boston.
- 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).

- Vasi, I. B., and B. G. King. 2012. Social movements, risk perceptions, and economic outcomes: The effect of primary and secondary stakeholder activism on firms' perceived environmental risk and financial performance. *American Sociological Review* 77(4):573–596.
- Vogel, D. 2010. The private regulation of global corporate conduct: Achievements and limitations. *Business & Society* 49(1) 68–87.
- von Krogh, G., J. Roos, and K. Slocum. 1994. An essay on corporate epistemology. *Strategic Management Journal* 15(1): 53–71.
- 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(9/10): 731–744.
- Weber, M. 1947. *The Theory of Social and Economic Organization*. New York: Free Press.
- Williamson, O. E. 1975. *Markets and Hierarchies: Analysis and Antitrust Implications*. New York: Free Press.
- Williamson, O. E. 2008. Outsourcing: Transaction cost economics and supply chain management. *Journal of Supply Chain Management* 44(2): 5-16.
- 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, Colombia: Universidad del Rosario.
- World Bank. 2013. Worldwide governance indicators: Frequently asked questions. <http://info.worldbank.org/governance/wgi/faq.htm>, accessed April 2013.
- 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(7): 469–486.
- Zandvliet, R. 2011. Corporate social responsibility reporting in the European Union: Towards a more univocal framework. *Columbia Journal of European Law* (online) 18: 38.

Appendix. Endogeneity Assessment

Our interviews with the social auditing firm that provided our data consistently indicated that auditors were assigned to audit teams according to three criteria: (1) their language skills, to ensure they could communicate with the audited supplier's managers and workers; (2) their availability, given their other auditing assignments; and (3) the need for at least one team member to qualify as a lead auditor. We conducted several empirical tests to assess whether lead auditors and non-lead auditors differed along dimensions that we hypothesized would affect the discovery and reporting of violations and found some evidence that they did:

- Graduate education. The distribution of educational attainment does not significantly differ between lead auditors and non-lead auditors, as indicated by a Pearson chi-squared test ($\chi^2 = 4.2$, $p = 0.24$) of an ordinal educational attainment variable coded 1 for high school, 2 for associate degree, 3 for bachelor's degree, and 4 for graduate degree.
- Gender. Lead auditors are no more likely than non-lead auditors to be a particular gender. Males make up 37% of the firm's lead auditors and 33% of its non-lead auditors, a non-significant difference according to a test-of-proportions analysis ($z = -1.06$, $p = 0.29$).
- Tenure. The average tenure of lead auditors is 4.4 years of service—significantly more than the 2.3 average for non-lead auditors (Wilcoxon rank-sum test: $z = -10.1$, $p < 0.01$).
- Audit skills training. Lead auditors averaged 5.5 audit skills training sessions—significantly more than the non-lead auditors' average of 2.6 (Wilcoxon rank-sum test: $z = -9.3$, $p < 0.01$).

The latter two results imply that larger audit teams would tend to have lower *average* tenure and lower *average* audit team skills training because the higher values for the lead auditors would be increasingly “diluted” by additional non-lead-auditor team members. However, in our analysis, we measure audit team tenure and audit skills training based on each team's *maximum* values, which do not suffer this “dilution” problem. (Furthermore, our model specification includes a series of dummies to control for audit team size.) These factors isolate our analysis from differences between lead and non-lead auditors, since all audit teams need one lead auditor.

We nonetheless conducted additional analyses to investigate whether endogeneity bias might affect our primary results. Given the differences in *audit skills training* and *tenure* between lead and non-lead auditors, we explored whether the estimated coefficients on our other hypothesized variables were substantially altered if we omitted those two variables from our model. The results of the more parsimonious model (reported in Column 2 of Table A-1) do not differ substantially from our primary results (reproduced in Column 1 of Table A-1). In particular, the coefficient magnitudes and statistical significance of the other hypothesized variables (*previous auditor*, *graduate education*, *all-female audit team*, and *mixed-gender audit team*) are remarkably stable across these two models. This indicates that irrespective of potential endogeneity concerns associated with *audit skills training* and *maximum tenure*, we find no evidence to suggest that such concerns spill over to the inferences associated with our other hypotheses (that is, H1, H3 when professionalism is measured by education, H4, and H5).

Our primary model is vulnerable to the possibility that endogeneity would bias our estimates on *maximum tenure* risk and maximum *audit skills training* if the auditing firm deployed teams whose members had higher maximum values of *audit skills training* and/or *maximum tenure* to particular types of establishment that varied in ways that were unobservable but would influence violations. One possible scenario is if the auditing firm assigned more highly skilled or experienced lead auditors to “bad apples”—that is, establishments suspected of being egregiously unconcerned with working conditions. Those lead auditors might be better able to ascertain information from managers and workers at such establishments, which are likely to have many violations. But the reverse might also be true: the auditing firm might send less-trained and less-skilled lead auditors to establishments suspected of having very safe working conditions; since such establishments would be expected to have few if any violations, less expertise would be required to audit them adequately. If these stories are true in our empirical context, we would expect to see much less variation in *audit skills training* and *maximum tenure* within the teams auditing the same establishment than between the teams auditing different establishments. In fact, we do not see such a pattern in our data when we decompose variation into within- and between-establishment components. For *audit skills training*, the within-establishment standard deviation (SD_w) is calculated based on all audit-level *audit skills training* values after de-meaning them at the establishment level and adding back the grand mean (that is, $x_{it} - \bar{x}_i + \bar{x}_i$). The between-establishment standard deviation (SD_b) is calculated based on establishment-level averages (that is, \bar{x}_i). For *audit skills training*, the between-establishment variation ($SD_b = 1.32$) is very similar to the within-establishment variation ($SD_w = 1.21$). That is, the variation in *audit skills training* among audit teams for two randomly drawn establishments is nearly identical to the variation in *audit skills training* among the audit teams of two randomly selected audits of the same establishment. The same is true for *maximum tenure*: the between-establishment variation ($SD_b = 1.51$) is very similar to the within-establishment variation ($SD_w = 1.44$). These results fail to support the notion that establishments tend to be consistently assigned teams with any particular average *audit skills training* or *maximum tenure*.

Establishments whose audit teams have the highest or lowest average *audit skills training* in our sample might be the most vulnerable to endogeneity, as they might represent the worst of the “bad apples” or the best of the “good apples.” We therefore reestimated our primary model on a subsample that excluded these outlier establishments whose audit teams’ average levels of *audit skills training* fell below the 1st percentile or exceeded the 99th percentile. Similarly, we also estimated our model on a subsample that excluded establishments whose audit teams’ average levels of *maximum tenure* fell below the 1st percentile or exceeded the 99th percentile. The results of these models, reported in Columns 3 and 4 of Table A-1, are nearly identical to the results of our primary model estimated on the full sample (Column 1). These results provide no evidence that endogeneity is driving our primary results.

Another approach to investigating whether endogeneity might be biasing our primary results is to instrument for the audit team’s maximum *audit skills training* and *maximum tenure*. We used average values of these characteristics among all auditors based in the auditing firm’s field office that staffed each establishment’s audits, an approach based on (a) Card’s (1995) instrumenting an individual’s propensity to attend college using the distance between that individual’s domicile and the nearest college and (b) Forman, Goldfarb, and Greenstein’s (2012) instrumenting a firm’s propensity to adopt Internet technology using the propensity of nearby firms. Because

these office averages vary little over time, the primary form of variation is cross-sectional (that is, between offices). We instrument for *auditing skills training*, *maximum tenure*, and *maximum tenure squared*—and also, to be as thorough as possible, *certification training*—by including the office-level corollaries to these variables using the audit year’s values at each audited establishments’ audit field office. The validity of our instruments requires that office-level averages of audit skills training and tenure: (1) be correlated with audit teams’ audit skills training and tenure (instrument relevance) and (2) be assumed not to have any direct influence on factories’ violation rates (instrument exogeneity). The first requirement is confirmed by observing that the specific audit-team-level and field-office-level variables are correlated at 0.77 for audit skills training, 0.58 for tenure, and 0.51 for certification training. The second condition relies on an assumption that field office demographics (our instrument) should not have a *direct* influence on audit results, but instead exert influence only via the demographics of the team members drawn from the office.

We sought to estimate this instrumental variables model using Poisson regression with endogenous regressors, but the matrix size created based on the 16,795 establishment-level fixed effects made this infeasible. As a second-best solution, we compared the results of our primary model estimated using fixed-effects OLS regression on the log number of violations (plus one to avoid losing cases with zero violations)—that is, assuming all variables were exogenous—to the results of a fixed-effects instrumental-variables OLS regression model on the logged number of violations (plus one). We used the log of the count as the dependent variable in these two models to make their specifications more comparable to those of our primary Poisson regression approach, which assumes that the logarithm of the violation count can be modeled by a linear combination of the independent variables. A Hausman test failed to reject the null, which is that the difference between the coefficients from the IV and OLS approaches is not systematic ($\chi^2 = 52.09$, $p = 0.16$). This implies that the IV approach in the continuous-dependent-variable context does not significantly alter the results, which provides no evidence suggesting that one cannot rely on the more straightforward modeling approach that assumes that the independent variables are exogenous. Given that (a) the OLS on the logged number of violations and (b) the Poisson model on the number of violations are each modeling the logged counts of violations, we infer that the IV Poisson model would not yield results systematically different from those of our primary fixed-effect Poisson model.

In sum, several alternative investigative approaches yield no evidence that endogeneity is biasing our results.

Table A-1. Regression results

Dependent variable: Number of violations

See notes for model definitions	(1)	(2)	(3)	(4)	(5)	(6)
Prior inspector	-0.042*	-0.039*	-0.040*	-0.043*	-0.044*	-0.034*
	[0.020]	[0.020]	[0.020]	[0.020]	[0.020]	[0.017]
Maximum tenure	0.065**		0.065**	0.073**	0.061**	0.049**
	[0.014]		[0.014]	[0.014]	[0.014]	[0.012]
Maximum tenure, squared	-0.004**		-0.004**	-0.005**	-0.004**	-0.003**
	[0.001]		[0.001]	[0.001]	[0.001]	[0.001]
Graduate education	0.028	0.027	0.012	0.027	0.026	-0.001
	[0.024]	[0.024]	[0.025]	[0.024]	[0.025]	[0.022]
Auditing skills training	0.020**		0.020**	0.019**	0.018**	0.024**
	[0.007]		[0.007]	[0.007]	[0.007]	[0.006]
All-female audit team	0.054**	0.057**	0.050**	0.052**	0.054**	0.041**
	[0.015]	[0.015]	[0.015]	[0.015]	[0.015]	[0.014]
Mixed-gender audit team	0.067**	0.075**	0.068**	0.065**	0.065**	0.064**
	[0.021]	[0.021]	[0.021]	[0.021]	[0.021]	[0.019]
Average age	-0.026	-0.011	-0.026	-0.027	-0.031+	-0.022
	[0.019]	[0.019]	[0.019]	[0.019]	[0.019]	[0.018]
Average age, squared	0.000	0.000	0.000	0.001+	0.001+	0.000
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Certification training	-0.024	0.021	-0.024	-0.026	-0.017	-0.030+
	[0.020]	[0.018]	[0.021]	[0.020]	[0.020]	[0.018]
Third-party protocol	-0.081	-0.084	-0.089	-0.077	-0.086	-0.112*
	[0.058]	[0.058]	[0.059]	[0.059]	[0.062]	[0.056]
Unannounced audit	0.029	0.028	0.034	0.025	0.025	0.057**
	[0.020]	[0.020]	[0.021]	[0.020]	[0.020]	[0.018]
Audit paid for by supplier or agent	-0.084**	-0.085**	-0.087**	-0.085**	-0.087**	-0.063*
	[0.026]	[0.026]	[0.027]	[0.026]	[0.026]	[0.025]
Re-audit	-0.348**	-0.348**	-0.341**	-0.349**	-0.358**	-0.297**
	[0.016]	[0.016]	[0.016]	[0.016]	[0.016]	[0.014]
Per-capita GDP (log)	-0.631*	-0.678**	-0.747**	-0.675*	-0.653*	-0.761**
	[0.262]	[0.262]	[0.265]	[0.264]	[0.263]	[0.243]
Regulatory quality	-0.177	-0.185	-0.262+	-0.163	-0.191	-0.221
	[0.150]	[0.150]	[0.157]	[0.151]	[0.152]	[0.141]
Press freedom	-0.511*	-0.530*	-0.532*	-0.576*	-0.559*	-0.464*
	[0.224]	[0.223]	[0.232]	[0.227]	[0.224]	[0.212]
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Number of auditors FE (2 to 5+)	Yes	Yes	Yes	Yes	Yes	Yes
Audit-sequence dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Client-country FE	Yes	Yes	Yes	Yes	Yes	Yes
Audits (N)	16,795	16,795	15,698	16,496	16,200	15,209
Firms	5,819	5,819	5,328	5,693	5,551	5,321

All results are from Poisson regression.

Brackets contain standard errors clustered by supplier; ** $p < 0.01$, * $p < 0.05$, + $p < 0.10$.

Column 1 reports estimates of the primary model (reproduced from Table 5, Column 1).

Column 2 reports estimates of a model that omits *maximum tenure* and *auditing skills training*.

Column 3 reports estimates on the subsample that excludes audited factories whose audit teams' average *auditing skills training* falls below the 1st percentile or exceeds the 99th percentile of the sample distribution.

Column 4 reports estimates on the subsample that excludes audited factories whose audit teams' average *maximum tenure* falls below the 1st percentile or exceeds the 99th percentile of the sample distribution.

Table 1. Geographic composition of supplier locations

	Audits		Factories	
	Number	Percent	Number	Percent
Africa	100	1%	38	1%
Americas	1,509	9%	522	9%
United States		949		285
Mexico		172		75
Brazil		84		37
Elsewhere in Americas		304		125
Asia and Australia	14,773	88%	5,084	87%
China (incl. Macao and Hong Kong)		11,746		3,917
India		708		277
Vietnam		424		153
Indonesia		377		137
Bangladesh		321		140
Philippines		270		96
Pakistan		184		71
Sri Lanka		159		61
Taiwan		131		56
Korea		120		49
Elsewhere in Asia & Australia		333		127
Europe	413	2%	175	3%
Turkey		186		72
Italy		88		42
Elsewhere in Europe		139		61
Total	16,795	100%	5,819	100%

Table 2. 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 3. 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 supplier or agent	0.56	0.50	0	1
Audit paid for by multinational firm	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 supplier or agent* and *audit paid for by multinational firm*, N =11,337 for *average age* and *maximum age*, and N =16,676 for *press freedom*.

Table 4. Pairwise correlations

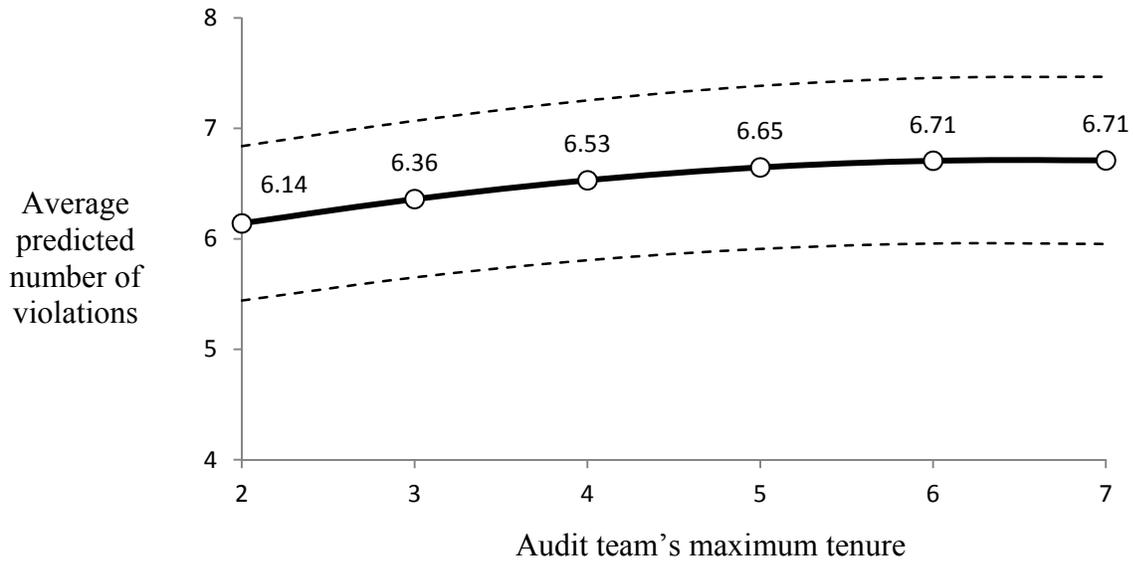
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)	
(1) Number of violations	1.00																						
(2) Previous auditor	-0.13	1.00																					
(3) Maximum tenure	-0.01	0.03	1.00																				
(4) Average tenure	-0.03	0.03	0.92	1.00																			
(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 supplier 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 multinational firm	-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	1.00

Table 5. Regression results
Dependent variable: Number of violations

	(1)	(2)	(3)	(4)	(5)	
	Coef.	Average marginal effects	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 supplier 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

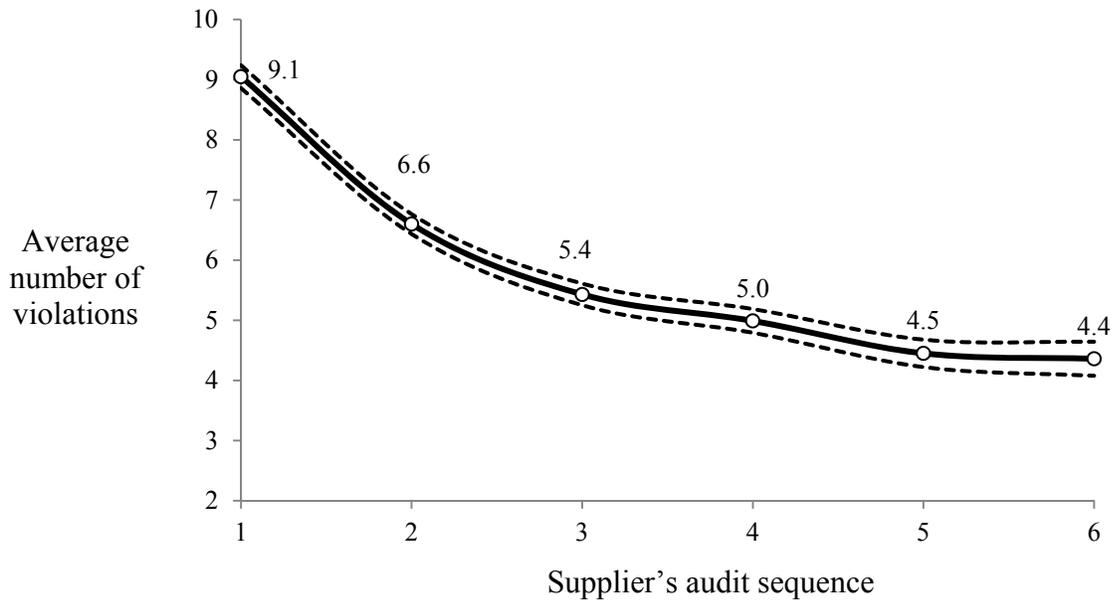
Standard errors clustered by supplier (factory); ** $p < 0.01$, * $p < 0.05$, + $p < 0.10$. All models also include fixed effects for the audited establishment, audit year, multinational firm country, number of auditors (2 through 5 or more), and the supplier's audit sequence (2nd through 6th or more). 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 supplier or agent* and *audit paid for by multinational firm* (983 audits), and *press freedom* (119 audits). Model 1 is the primary model estimated on the entire sample. Model 2 excludes audits conducted for multinational firms 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 audit during the sample period.

Figure 1. Effect of audit team’s maximum tenure on average predicted violations per supplier



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

Figure 2. Decline of average number of violations per audit in successive audits



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