Beyond Symbolic Responses to Private Politics:
Codes of Conduct and Improvement in Global Supply Chain Working Conditions

Jodi L. Short
UC Hastings College of the Law

Michael W. Toffel
Harvard Business School

Andrea R. Hugill
Georgetown University McDonough School of Business

June 15, 2018

Worker rights advocates seeking to improve labor conditions in global supply chains have engaged in private politics that led transnational corporations (TNCs) to adopt codes of conduct and to monitor their suppliers for compliance, but it is not clear whether or when these organizational structures can actually raise labor standards. We extend the literature on private politics and decoupling by identifying structural contingencies in the institutional environment and in program design under which codes and monitoring are more likely to go beyond mere symbolism and to be associated with improvements in supply chain working conditions. At the institutional level, we find that suppliers improve working conditions more when they face greater exposure risk from their domestic civil society, and when their buyers are more sensitive to such exposure. At the program design level, we find that suppliers improve more when the monitoring regime signals a cooperative approach and when auditors are highly trained. We also identify several structural contingencies among and between institutional and program design features. These findings advance theory and provide new empirical insights on the outcomes of private political activism and suggest key considerations to inform monitoring strategies aimed at improving working conditions in global supply chains.

Since the 1990s, private political strategies have become increasingly vital to the fight for better working conditions in global supply chains. While states remain key actors in promulgating and enforcing labor standards, the "anti-sweatshop movement arguably made the TNC [transnational corporation] into the central locus of struggle over labor rights and globalization at the turn of the twenty-first century" (Bartley and Child, 2014: 657). Activists and consumers have mobilized highly publicized private politics campaigns against corporate targets like Nike (for child labor in its supplier factories), Apple (in response to worker suicides

and reports of rampant labor abuses at its Foxconn factory), and prominent European apparel brands (following the catastrophic collapse of the Rana Plaza building where many of their garments were produced, killing more than 1,000 Bangladeshi factory workers). These targeted companies suffered reputational damage and incurred financial costs to address harms to workers.

Beyond high-profile cases like these, private political activism has had far-reaching effects in shaping the way TNCs structure the governance of their global value chains. To avoid the negative publicity generated by anti-sweatshop activism, thousands of TNCs, including all US *Fortune* 500 companies, have adopted codes of conduct that require suppliers to meet specified workplace standards (McBarnet, 2007). Many also conduct "social audits" to monitor their global value chains and assess suppliers' adherence to those codes (Short, Toffel, and Hugill, 2016). Codes and monitoring are also used by the scores of multi-stakeholder initiatives that provide collective fora for private regulation of supply chain practices. Organizations like the Roundtable on Sustainable Palm Oil, the Electronics Industry Citizenship Coalition, and the Ethical Trading Initiative bring together representatives from industry and civil society to set standards for supply chain practices and certify compliance through social auditing. Hundreds of thousands of audits are conducted on behalf of individual firms and multi-stakeholder initiatives each year (Gould, 2005), making social auditing itself an $80 billion industry (AFL-CIO, 2013).

While TNCs initially adopted codes and monitoring for business purposes, including information gathering and reputation management, these organizational structures have been embraced by nongovernmental organizations (NGOs) and other international organizations as part of a broader private political strategy to improve conditions for global supply chain workers (Utting, 2005; LeBaron and Lister, 2015). Codes and monitoring are central to the strategy of

NGOs that alternately agitate against and partner with TNCs to encourage ever-stricter standards and more robust private monitoring (Bartley, 2007; King and Pearce, 2010; LeBaron and Lister, 2015). Codes and monitoring are the lynchpin of voluntary United Nations initiatives in the area of business and human rights, like the Global Compact and the Guiding Principles on Business and Human Rights, which rely on companies to implement private standards and monitoring to improve global labor and other human rights conditions (Ruggie, 2008). Responding to these calls for improvement, leading social auditing firms advertise that their monitoring services will help "both suppliers and customers in implementing sustainable business practices and *improving* workplace conditions in global supply chains [emphasis added]" (UL Responsible Sourcing, 2015; see also Elevate, 2016; Intertek, 2016).

However, it is not clear that codes and monitoring are associated with meaningful improvements in supplier labor practices. Many have argued that these organizational structures are, at best, window-dressing (Esbenshade, 2004; Frynas, 2005; Seidman, 2007; Barkemeyer, 2009) or, worse, a calculated ploy to undermine labor organizing efforts (Justice, 2006) and fend off more stringent state regulation (Utting, 2005; Barkemeyer, 2009; Shamir, 2011). Certainly, codes and monitoring bear all the hallmarks of "myth and ceremony" (Meyer and Rowan, 1977). It would be difficult to find organizational structures that more "dramatically reflect the myths of their institutional environments instead of the demands of their [organizations'] work activities" (Meyer and Rowan, 1977). Companies adopted codes and monitoring largely in response to coercive and mimetic external pressures (Suchman, 1995). Initially, they adopted these self-regulatory regimes for business purposes: to deflect or avoid negative publicity in order to preserve assets like brand value and reputation (Bartley and Child, 2014; Mena and Waeger, 2014). Codes and monitoring diffused throughout the field of TNCs as companies imitated the

aggressively marketed embrace of these structures by industry leaders like Nike and Apple and as stakeholders like NGOs and shareholder activists made them touchstones of transnational corporate legitimacy (Etzion and Ferraro, 2010). However, the substantive changes in production practices that such structures formally demand—for instance, observing minimum wage requirements, overtime restrictions, and freedom of association rights for workers—are fundamentally at odds with the economic logic of the extended global value chain (Locke, 2013). Thus, these structures are ripe for decoupling (Meyer and Rowan, 1977; Meyer et al., 1997).

Private political activists nevertheless continue to demand adoption of these structures as a part of their strategies to change suppliers' labor practices. To date, private politics scholarship has paid little attention to the link between organizational structures adopted in response to activism and changes in organizational practices, focusing instead on the symbolic responses triggered by corporate-targeted activism (McDonnell and King 2013). While those are important, it is vital to look beyond symbolism and investigate the relationship between activism-provoked organizational structures and changes (if any) in the harmful practices they were meant to address. We argue that decoupling is not inevitable, despite the classic markers that would predict it in this context. We extend the literatures on private politics and decoupling by investigating the structural contingencies that account for variation in the extent to which supply chain codes and monitoring are associated with improvement in working conditions. Specifically, we investigate how compliance improvement rates differ among suppliers depending on both institutional compliance pressures and the design of monitoring programs. We also, for the first time, examine the interaction of institutional pressures and program design as well as the interaction among program design elements.

Using a novel dataset drawn from thousands of audits for code-of-conduct compliance in 66 countries by one of the world's largest supply chain auditing firms, we find that suppliers improve more when (a) local and global institutional pressures create risks that harms to workers will be exposed, (b) monitoring programs are designed to signal a cooperative approach, and (c) auditors are highly trained. We also find that the improvement associated with greater auditor training is enhanced in the context of a cooperation-signaling audit program design—more so than when suppliers face higher levels of exposure risk.

By identifying structural contingencies which favor coupling, these findings challenge the assumption that private politics inevitably produce organizational structures decoupled from practices. Our results also advance theory and empirical research on the organizational outcomes of private political activism for social change and suggest key considerations that should inform the design and implementation of monitoring strategies to improve conditions in global supply chains.

PRIVATE POLITICAL ACTIVISM AND DECOUPLING

Private political strategies seek to change the behavior of private market actors like corporations by mobilizing activism to target those actors directly (Baron, 2003; King and Pearce, 2010). An extensive private politics literature has focused on activists' targeting strategies (Bartley and Child, 2014; McDonnell, King, and Soule, 2015), companies' motivations for submitting to activists' demands (King and Soule, 2007; King, 2008), and the organizational structures corporate targets have adopted in response to private political activism. Studies reveal, for instance, that that private political activism has prompted firms to adopt "impression management tactics" (McDonnell and King, 2013: 411), public "concessions" to conform to activists' demands (Eesley and Lenox, 2006; King, 2008), policies sanctioning certain brands

targeted by activists (Briscoe, Gupta, and Anner, 2015), and corporate social responsibility (CSR) officer positions or board committees (McDonnell, King, and Soule, 2015). However, private politics research has not sought to determine whether these organizational structures are related to positive changes in organizational behavior that align it more closely with activists' normative goals. Instead, private politics scholars have suggested that symbolic outcomes are the appropriate focus of research. Arguing that markets "are inherently grounded in symbolic understandings about what is valuable and appropriate" (McDonnell and King 2013: 410), they contend that changing these understandings can promote the market transformations sought through private politics.

We suspect that this focus is at least partly an artifact of the decoupling literature, which strongly suggests that organizational structures adopted in response to corporate-targeted activism are likely to be symbolic, or decoupled from organizational activities. First, there is a strong consensus in the literature that organizational structures adopted to gain legitimacy with external stakeholders rather than to satisfy the task-related efficiency demands of production will tend to be implemented symbolically and decoupled from practices (Meyer and Rowan, 1977; Boiral, 2007; Bromley and Powell, 2012). Second, symbolic structures are more likely to be decoupled in contexts where efficiency demands are strong and not tempered by countervailing institutional pressures (Meyer and Rowan, 1977; Meyer et al., 1997). Finally, resource constraints impede substantive implementation of formal organizational structures (Meyer et al., 1997; Bromley and Powell, 2012; Lim and Tsutsui, 2012). Suppliers to global value chains are subject to all three of these constraints: they face intense efficiency demands to produce high volume at low cost (Gereffi et al., 2005) and codes of conduct threaten to raise the cost of labor, a key source of competitive advantage; many are in countries with weak regulatory institutions

and lax enforcement of labor standards; and many lack the resources to effectively implement formal structures like codes and monitoring. Thus, the axioms that emerge from the decoupling literature suggest that organizational structures like codes and monitoring are likely to be ceremonial window dressing "implemented, evaluated, and monitored so weakly that they do little to alter daily work routines" (Meyer and Rowan, 1977; Bromley and Powell, 2012; Lim and Tsutsui, 2012) in ways that might improve conditions for workers.

While research in this domain has mostly theorized and documented decoupling, a growing stream focuses on the conditions under which organizational structures adopted symbolically are actually implemented substantively or become coupled with organizational practices (Bromley and Powell 2012; Bartley and Egels-Zandén, 2015). Consistent with the decoupling literature, most studies that do find coupling attribute it to coercive institutional pressures, particularly to forms of state power, such as regulatory inspection and enforcement (Dobbin and Kelly, 2007; Short and Toffel, 2010; Marquis and Qian, 2014). Other studies identifying successful coupling of symbolic structures have been of voluntary programs implemented in the context of broader, legally backed state regulatory regimes such as US antidiscrimination law (Kalev, Dobbin, and Kelly, 2006) or environmental law (Potoski and Prakash, 2005). Similarly, studies specifically investigating suppliers' compliance with labor codes of conduct have found that codes and monitoring tend to be associated with better working conditions when combined with government regulatory efforts (e.g., Rodríguez-Garavito, 2005; Seidman, 2007; Amengual, 2010; Locke, Rissing, and Pal, 2013; Toffel, Short, and Ouellet, 2015; Amengual and Chirot 2016). In addition to the coercive power of the state, studies have found that institutional pressures from civil society actors like unions (Bartley and Egels-Zandén, 2015), a free press (Toffel et al., 2015), NGOs (Seidman, 2007; Fransen, 2012; Zajak, 2017), and

brands (Oka, 2010a; Bartley and Egels-Zandén, 2015) can induce suppliers to couple their symbolic commitments to codes and their labor practices.

Recent studies have expanded the decoupling literature's traditional focus on coercive institutions to investigate how the activities of individual organizational actors can create contingencies that promote coupling (Espeland, 1998; Hallett, 2010; Overdevest, 2010; Tilcsik, 2010; Bartley and Egels-Zandén, 2015). For example, Tilcsik (2010) demonstrates how personnel turnover and changes in intra-organizational power dynamics led a government agency to fully implement a budgeting system that it had initially adopted symbolically to appease external stakeholders. In a study of Indonesian apparel and footwear factories, Bartley and Egels-Zandén (2015) find that the coupling of labor codes of conduct and supplier labor practices was contingent on local union members' ability to leverage ties with brands, international NGOs, and global unions to pressure suppliers to live up to their commitments to codes of conduct.

Collectively, the literature on coupling and decoupling contains important insights into the challenges and contingencies of coupling organizational structure and practice, but it also contains gaps that have hindered dialogue with the private politics literature. First, the decoupling literature's traditional focus on the coupling power of state-based and other coercive institutional pressures has limited its ability to explain variation in substantive outcomes observed among corporate targets of private political activism that adopt organizational structures in environments where coercive institutions are weak or lacking. But these are precisely the types of environments in which private political campaigns operate. Thus, coupling research must address coupling contingencies outside of the traditional channels if it is to tackle the contexts in which private political strategies are most likely to be deployed.

Second, while qualitative coupling studies have begun to address this concern by revealing how individual actions within organizations can promote coupling even under challenging institutional conditions, their focus on the "thorny, on-the-ground processes" (Bartley and Egels-Zandén, 2015: 3) in a single firm (Hallett, 2010; Tilcsik, 2010) or in a few firms in the same institutional context (e.g., Bartley and Egels-Zandén, 2015) neglects structural contingencies of coupling. Structural contingencies are particularly important to understanding the outcomes of private political activism and developing private political strategies, because it is difficult to predict or evaluate the success of private politics if success depends on local, individual actions that may be idiosyncratic. Coupling research can be made more relevant to private politics if it expands the range and types of coupling contingencies it investigates and the methodologies it uses to identify them.

Third, there has been insufficient focus on the relationship between symbolic structures and organizational *change* in the decoupling literature. Many prominent, large-scale decoupling studies measure outcome variables at one point in time (e.g., Westphal and Zajac, 2001; Zajac and Westphal, 2004; Lim and Tsutsui, 2012; Marquis and Qian, 2014; Toffel et al., 2015). While such measures can determine whether formal organizational structures are credible signals of organizational practices, they reveal little about whether such structures are associated with *changes* in organizational practices, which is, after all, the explicit goal of private political activism (Baron, 2003; King and Pearce, 2010).

In our empirical context, no existing study addresses the precise question of why some suppliers improve their compliance with codes of conduct more than others. Most research in this field has measured levels of supplier compliance at a particular moment in time rather than compliance improvement over time (e.g., Egels-Zandén, 2007; Locke, Qin, and Brause, 2007;

Oka, 2010a, 2010b; Ang et al., 2012; Bartley and Egels-Zandén, 2015; Toffel, Short, and Ouellet, 2015). Thus, these studies cannot address questions about the conditions producing changes in organizations' market behaviors. A few studies have observed that, in the aggregate, supplier compliance with codes has improved over time (Locke, Qin, and Brause, 2007; Shea, Nakayama, and Heymann, 2010; Nadvi et al., 2011; Ang et al., 2012; Locke, Rissing, and Pal, 2013; Toffel, Short, and Ouellet, 2015). Others have shown that labor code compliance tends to improve more rapidly in some code categories, like health and safety, than in others, like freedom of association (Barrientos and Smith, 2007; Ruwanpura, 2012). However, these studies have not hypothesized what conditions are associated with improvement.

The few studies that have empirically examined factors associated with code compliance improvement have significant limitations. For instance, Weil and Mallo (2007) and Ang et al. (2012) investigated factors associated with improvements in labor standards compliance, but both studies measured these factors, as well as the improvement rates themselves, using data aggregated by market (that is, at the national or regional level), leaving open the question of what accounts for individual suppliers' improvement (or failure to improve). Qualitative studies discussing factors associated with observed improvements in code compliance (e.g., Egels-Zandén, 2007; Locke, Rissing, and Pal 2013) similarly fail to identify factors that can explain why some factories improve while others do not, because all of the factories in their small samples improved. To our knowledge, no study has investigated the factors associated with variation in suppliers' rates of improvement.

We seek to fill these gaps and promote more meaningful dialogue between the private politics and decoupling literatures by attending to the structural contingencies associated with the coupling of labor codes of conduct and supplier labor practices. Our hypotheses look beyond

state-based institutions to investigate the coupling potential of structural contingencies operating at two levels: institutional and programmatic. Specifically, we hypothesize how the coupling of labor codes of conduct and supplier labor practices is likely to be associated with ongoing institutional pressure from civil society groups on suppliers and brands, as well as key design features of the monitoring programs that targeted brands are requested to adopt. We also investigate how institutional pressures and program design create contingencies for each another.

We selected these structural contingencies for two reasons. First, the institutional pressures and program design features we hypothesize are conditions that private political activists can influence. For instance, activists pressuring firms to adopt codes of conduct and supplier monitoring can insist that these programs contain certain elements and can mobilize other civil society actors to pressure firms to comply. Thus, these structural contingencies are particularly relevant to understanding and strategizing the outcomes of private politics. Second, with few exceptions (e.g., Kalev, Dobbin, and Kelly, 2006), little attention has been given to the influence of program design on coupling outcomes and, to our knowledge, none has been paid to the interaction between program design and institutional conditions. We suspect that this is because the decoupling literature tends to view programs adopted in response to activism as mere myth and ceremony. We argue that this perspective misses important opportunities to make them more meaningful, and we take seriously the contribution of program design to coupling outcomes.

Institutional Compliance Pressures Created by Private Political Activism

In their fight to improve supply chain labor conditions, "[a]ctivists' main weapon against corporations is their ability to threaten corporate reputations by exposing malfeasance, lack of ethical decision-making, or the use of normatively questionable practices" (King, 2014: 203).

Global buyers are particularly sensitive to the possibility that their suppliers' labor abuses will be exposed, because even one incident in the supply chain can damage carefully cultivated corporate reputations (Oka, 2010a). Buyers thus face an aggregated reputation risk that well exceeds the probability of any single supplier being exposed. For example, a buyer with 100 suppliers, each with a 1-percent chance that its poor labor practices will be exposed in a given year, faces a 63-percent chance that at least one of its suppliers will be exposed that year (calculated as 1-(1-0.01)^100). This aggregated risk encourages buyers to pressure suppliers to improve. Moreover, research has long demonstrated that it is the *perception* of exposure risk rather than the actuarial probability of exposure that influences behavior (Gibbs, 1975). Thus, even when the probability of detection is small, it may deter wrongdoing if corporate decision-makers perceive the risk of exposure to be significant. For instance, a recent study found that foreign direct investment is often deterred because managers have highly inflated perceptions of the political risks to capital in emerging economies (Giambona et al., 2017).

Below, we explore two distinct but related dimensions of exposure risk faced by global buyers: the potential for supplier abuses to be exposed and buyers' sensitivity to the possibility of exposure. We argue that both will prompt buyers to pressure suppliers to improve and thus will be positively associated with supplier improvement.

Exposure risk. Although government inspection regimes are often weak in the countries where global suppliers are located, research suggests that civil society actors like NGOs and the press can provide monitoring and expose wrongdoing (Seidman, 2007; Mattli and Woods, 2009; Fransen, 2012; Zajak, 2017). The local press and local NGOs play symbiotic roles in transnational advocacy networks that promote global norms such as labor standards and human rights (Keck and Sikkink, 1998: 3). The high-profile, international NGOs that are often at the

center of such networks depend heavily on local NGOs to collect information about violations of global norms by local actors. In fact, some recent studies have found local organizations to be more important than their global counterparts in transnational advocacy campaigns (Zajak, 2017). The strategy of the global anti-sweatshop movement has been to work with local NGO partners to identify which local firms supply targeted global brands and to do the painstaking investigative work required to reveal exploitive labor practices at these suppliers (Bartley and Child, 2014; Zajak, 2017). Local NGOs, in turn, depend on domestic media and domestic channels of communication to make exploitive practices known to their more powerful international counterparts in the advocacy network (Keck and Sikkink, 1998; Bartley and Child, 2014; Zajak 2017). The more free and open these information channels, the more likely local abuses are to attract local and ultimately international attention, condemnation, and discipline (King, 2014).

To date, research on civil society pressure and standards compliance has focused on how monitoring by civil society actors affects compliance *levels* at a particular point in time. Toffel, Short, and Ouellet (2015), for instance, have demonstrated that suppliers in countries with more press freedom exhibit higher levels of compliance with codes of conduct. However, it is not clear that the mechanisms fostering high compliance *levels* will also foster *improvements* in compliance. For instance, in high-compliance environments, there may be less room to improve due to a lack of low-hanging fruit (Chatterji and Toffel, 2010) or simply less pressure to improve. Similarly, it is not clear that the presence of civil society actors like NGOs will be related to compliance improvement in the same way that it relates to compliance levels. For instance, domestic environments with very low or very high levels of labor standards compliance might attract more NGOs, but it is a distinct and important question whether NGOs will go on to

contribute to *improvement* in labor practices.

We argue that institutional pressure generated by civil society actors like the press and NGOs will be associated with improvements in supplier compliance. When suppliers' failure to meet the global norms prescribed by codes of conduct is documented in social audits, those suppliers become attractive targets for transnational advocacy networks seeking to raise international labor and human rights standards.[1] These networks' *modus operandi* is to identify violators of global norms and induce them to change (Keck and Sikkink, 1998: 3). Local NGOs in these networks see their role as "promoting change by reporting facts" that can attract the attention and support of international NGOs, press, and policymakers (Keck and Sikkink, 1998: 19). A free local press helps them discover and publicize facts about labor abuses at suppliers, increasing the prospect that such suppliers will be disciplined by their buyers or will suffer domestic political, legal, or economic consequences (Fransen, 2012; Berliner et al., 2015; Zajak, 2017).

We therefore hypothesize:

> Hypothesis 1 (H1): Suppliers will improve more when located in institutional environments in which there is greater potential for civil society monitoring mechanisms to expose noncompliance with codes of conduct.

Buyer exposure sensitivity. The institutional pressure generated by civil society

[1] In a context very different from ours, studies have suggested that activists and the press are less likely to target the worst-behaved companies than companies with strong reputations for social responsibility performance (Luo, Meier, and Oberholzer-Gee, 2012; Bartley and Child, 2014) or companies that have adopted extensive organizational structures to implement their CSR initiatives (McDonnell, King, and Soule, 2015). We do not dispute these findings, but we note that they focus on very different contexts and actors than those we study here. Those studies focused on activism directed toward branded multinational companies with reputations to protect and argued that activists can exercise more leverage over such firms: the better a brand's reputation, the more significant the financial consequences of reputational damage. Such is not the case with the subjects of our study: local suppliers in developing countries. They are unbranded, largely invisible to consumers, and thus more insulated from this kind of reputational threat. As suggested in the literature on transnational advocacy networks, we believe that activists will select their targets very differently in these local settings and will attempt to identify the worst practices by local suppliers in order to gain the most leverage over the global brands that are their ultimate targets.

activism depends not only on the probability that such activism will bring certain suppliers' wrongdoing to light, but also on how buyers working with those suppliers are likely to react. The reputational stakes of exposure are higher for some buyers than for others. For instance, buyers with particularly high-value reputations might be acutely sensitive to negative publicity (Abito, Besanko, and Diermeier, 2016). Indeed, in the supply chain labor context, research has demonstrated that buyers that are highly reputation-conscious are more likely to work with suppliers that better comply with labor standards (Oka, 2010a).

Alternatively, buyers might be especially sensitive to negative publicity if their reputations have already been tarnished. Economic models predict that such companies are more likely to invest in self-regulatory measures to avoid additional reputational harm (Abito, Besanko, and Diermeier, 2016). In addition, a damaged reputation can invite more activism, further incentivizing protective measures (Abito, Besanko, and Diermeier, 2016; Dorobantu, Henisz, and Nartey, 2017). Empirical studies confirm that companies with more reputational damage in the past are more likely to take actions to protect their reputation (Kotchen and Moon, 2012; McDonnel and King, 2013; McDonnell, King, and Soule, 2015). What these studies do not reveal, however, is whether such protective measures reduce the social harms that gave rise to the reputational harm. We argue that they will, because reputation-compromised buyers face greater reputational risk if additional information about their suppliers' harmful practices is exposed, and they are therefore more likely to try to ensure that their suppliers correct abuses.

> Hypothesis 2 (H2): Suppliers will improve more when they produce for buyers that have been publicly exposed for harms to workers in their supply chain.

Program Design Features

A monitoring system has numerous elements, from the frequency and rigor of inspections to the composition of the inspection team. Design choices about what elements to include will be

driven by the goals of monitoring. We start from the premise that if improvement, rather than mere surveillance, is an objective, the monitoring regime must be designed and implemented in ways that encourage the transfer of compliance-related knowledge from auditors to suppliers. Below, we identify design features that might facilitate such exchange.

Signaling a cooperative approach to social auditing. One important design choice is whether audits will be announced in advance. Some argue that this gives suppliers time to cover up their bad behavior (Clean Clothes Campaign, 2005; AFL-CIO, 2013; LeBaron and Lister, 2015) and there is empirical evidence for this (Gray, 2006; Marks, 2012; Toffel, Short, and Ouellet, 2015). Worker rights advocates therefore have long favored unannounced audits (Frenkel and Scott, 2002). Others, however, maintain that unannounced audits meant "to catch managers unaware . . . may, for now, help buyers to learn more about the real conditions in factories," but this "aggravates the relationship between buyers and suppliers ... [and] make[s] it difficult to achieve any sustainable change" (Gould, 2005: 28).

If the goal of an audit is to discover as much as possible about supplier conditions and suspected wrongdoing, evidence suggests that unannounced audits will be more effective (Gray, 2006; Marks, 2012; Toffel, Short, and Ouellet, 2015). But if the goal is to foster supplier improvement, the choice is less clear. It is plausible that suppliers might be motivated to improve their practices if they know that their wrongdoing can be discovered and subject to punishment at any time through a surprise audit. However, most buyers do not conduct audits regularly enough for this to be a serious threat. Moreover, studies have suggested that a punitive, "policing"-style approach to monitoring can undermine firms' motivation to cooperate with regulators (Bardach and Kagan, 1982; Ayres and Braithwaite, 1992; Short and Toffel, 2010).

Consequently, a consensus has begun to emerge among academics and practitioners that

suppliers are more likely to improve with a more cooperative or "commitment-oriented" approach to monitoring (Locke, Amengual, and Mangla, 2009). Rather than using audits to detect violations and threaten sanctions, the cooperative approach provides an opportunity "to engage in a process of root-cause analysis, joint problem solving, information sharing, and the diffusion of best practices that is in the mutual self-interest of the supplier, the auditors, and the global corporations for which they work" (Locke, Amengual, and Mangla, 2009: 321). The underlying theory, developed most extensively by Ayres and Braithwaite in *Responsive Regulation* (1992), is that regulators' cooperation will be reciprocated with compliance.

Studies have suggested that a cooperative approach to monitoring can help buyers, suppliers, and auditors develop trusting relationships that are more likely than punitive, arms-length approaches to improve compliance (Frenkel and Scott, 2002; Locke and Romis, 2007). Locke and Romis (2007), for instance, find better compliance with labor standards at a Nike supplier that had regular, face-to-face contact with Nike's compliance staff than a supplier that had a more distant, formal relationship with Nike's monitors. They argue that these ongoing interactions led to "transparency and trust … as well as joint problem solving" (Locke and Romis, 2007: 60). By contrast, the plant that had less direct interaction viewed the company's standards as little more than hurdles to clear in order to continue receiving orders.

Although we do not observe the micro-level interactions among the buyers, suppliers, and auditors in our sample and so cannot speak to the qualitative aspects of these relationships, we argue that buyers signal trust and a willingness to adopt a cooperative approach to monitoring when they give suppliers advance notice of audits. At the very least, this indicates trust in the formal sense of making the buyer vulnerable to the possibility of opportunism on the part of suppliers who might use the time afforded by advance notice to hide their misdeeds (Mayer,

Davis, and Schoorman, 1995). It is this very vulnerability that makes announcing audits so controversial. Our interviews with brands' ethical supply chain managers and with social auditors employed by several auditing firms consistently indicated that unannounced audits convey distrust and a punitive or "policing" approach to monitoring, with auditors sometimes denied entry to factories, whereas announced audits convey a more trusting and cooperative approach. Thus, although an announced audit might uncover fewer violations, we argue that it can pay dividends down the road by encouraging suppliers to reciprocate with increased compliance.

> Hypothesis 3 (H3): Suppliers that receive advance notice of audits will improve more than other suppliers.

Auditor training. There is much skepticism about whether social auditing can foster improvement and questions have been raised about the competence of the auditors and the integrity of the auditing process (O'Rourke, 2002; Esbenshade, 2004; LeBaron and Lister, 2015). Critics charge that auditors lack the knowledge and independence to detect labor abuses (O'Rourke, 2002; Esbenshade, 2004; Locke, Amengual, and Mangla, 2009; AFL-CIO, 2013), that they "shade their findings depending on the client" (LeBaron and Lister, 2015), that they are easily duped by managers who cook the books and coach employees to lie about conditions (AFL-CIO, 2013), and that some are outright corrupt (Clean Clothes Campaign, 2005). Others see auditors stuck in an ever-more-sophisticated "cat and mouse" game with suppliers that maintain fake wage and hours records and coach their workers on how to answer auditors' questions (Power, Ng, and Singh, 2008; Karunakaran, 2013; LeBaron and Lister, 2015). Research has indeed documented biases in the way social auditors record violations based, for instance, on whether or not the supplier is paying for the audit and on whether or not the auditor and supplier have a longstanding relationship (Short and Toffel, 2016; Short, Toffel, and Hugill,

2016). Some, including critics of social auditing, have suggested that auditors can be more effective when they are more highly trained (Locke, Amengual, and Mangla, 2009; AFL-CIO, 2013) and research has found that better-trained auditors identify more violations than less-trained ones (Short, Toffel, and Hugill, 2016).

We argue here that training will likewise enable auditors to help suppliers improve following an audit. Studies have documented that social auditors play an important "pedagogical" role, often instructing factory managers how to remedy the violations they find (Amengual, 2010). Our interviews with auditors and managers at social auditing firms reveal that the auditors' training typically teaches them how to find violations and what conditions tend to cause them. Such training probably makes auditors better able to identify root causes and help suppliers devise solutions. Recent evidence has indicated that government inspections can prompt improved working conditions (Levine, Toffel, and Johnson, 2012), suggesting that inspectors might play a dual role of assessing conditions and suggesting how to improve. Indeed, studies in the knowledge transfer and learning literatures find that certain types of training can improve the ability to apply and convey information in personal interactions (Thompson, Gentner, and Loewenstein, 2000; Loewenstein, Thompson, and Gentner, 2003; Nadler, Thompson, and Van Boven, 2003) and that information is more likely to be absorbed and acted upon when it comes from a source perceived to have expertise (Borgatti and Cross, 2003; Thomas-Hunt, Ogden, and Neale, 2003; Reinholt, Pederson, and Foss, 2011). We therefore hypothesize:

> Hypothesis 4 (H4): Suppliers will improve more following audits conducted by audit teams that are more highly trained.

Interaction of Institutional Pressures and Program Design Features

Auditor training in the context of a cooperative approach to auditing. Improving

supplier practices through auditing requires not only a knowledgeable auditor, but also a supplier willing and able to receive the information. A substantial literature suggests that individuals and organizations share and absorb knowledge more effectively in collaborative, cooperative, and trusting relationships (Coleman, 1988; Szulanski, 1996; Dyer and Chu, 2003; McEvily, Perrone, and Zaheer, 2003; Inkpen and Tsang, 2005; Obstfeld, 2005). For instance, Cheng, Yeh, and Tu (2008) show that the transfer of green production practices from buyers to suppliers is most effective when buyers let suppliers participate in decision making *and* when those buyers and suppliers trust one another. Buyers and suppliers surveyed by Oka (2010b) similarly reported more learning about compliance with workplace standards in trusting relationships. As we argue above, announcing audits signals trust in the supplier; we therefore expect suppliers who receive advance notice to be more receptive to the knowledge auditors have to convey in those audits. Because highly trained auditors likely will have more and/or higher-quality information to convey, we hypothesize the following moderated relationship:

> Hypothesis 5 (H5): Suppliers audited by highly trained auditors will improve more following announced audits than following unannounced audits.

Auditor training and institutional compliance pressure. In contrast to demonstrated synergies between learning and cooperation, studies have found that the threat of punishment can be detrimental to learning (Deci and Ryan, 1985), including learning how to improve compliance (e.g., Bardach and Kagan, 1982; Ayres and Braithwaite, 1992). Tenbrunsel and Messick (1999: 688) explain that fear of punitive sanctions can blunt improvement because it renders the decision to comply "mainly about averting penalties or achieving rewards" rather than about achieving the best performance possible. Firms motivated primarily by the fear of sanctions may learn enough to avoid the sanctions, but may not be motivated to improve any further.

In addition, a punitive approach to enforcement can sour relationships between suppliers

and auditors, making it "very hard to initiate any improvements" (Gould, 2005: 28). Researchers have suggested that suppliers subjected to punitive monitoring are more likely to try to hide their wrongdoing, thus shifting scarce resources away from the task of actually improving working conditions (Plambeck and Taylor, 2016). In a different context, Short and Toffel (2010) demonstrated that firms threatened with penalties if they failed to adopt an internal compliance program did not implement it successfully and did not improve compliance.

The foregoing arguments suggest that fear and reciprocity are two possible and potentially competing motivations for suppliers to improve compliance after an audit. Those facing external institutional compliance pressures may fear detection or damaging publicity. Those who receive signals of trust through the design and implementation of their monitoring regime may desire to reciprocate that trust. We argue that highly trained auditors will foster more learning and consequently more improvement amongst suppliers motivated to reciprocate buyers' trust signals (such as advanced notice of audits) than amongst suppliers primarily motivated by fear, which in our context refers to suppliers at greater risk of having worker abuses exposed and who serve buyers more sensitive to such exposure.

> Hypothesis 6a (H6a): Auditor training will have a greater influence on the improvement of suppliers given advance notice of audits than it will on the improvement of suppliers at greater risk of exposure by civil society monitoring mechanisms.

> Hypothesis 6b (H6b): Auditor training will have a greater influence on the improvement of suppliers given advance notice of audits than it will on the improvement of suppliers whose buyers are more sensitive to exposure risk.

DATA AND METHOD

Empirical Context and Sample

We tested our hypotheses using data from code-of-conduct audits conducted by a large social auditing company (henceforth, the "social auditor") that requested anonymity. The data

include audits conducted from 2004 through 2009, the most recent six-year period for which we could obtain access. Various characteristics of the audits, auditors, and audited suppliers were provided, including unique identifiers (but not names) for the auditors, the suppliers, and the buyers on whose behalf the audits were conducted. While many buyers issue their own supplier codes of conduct, our discussions with the social auditor revealed that the differences between these codes are slight, which gave us confidence in treating all of these audits similarly.

Because our empirical specification requires data from a supplier's focal (current) audit and its prior audit, our sample is limited to those suppliers for which our data includes at least two audits. Our estimation sample consists of 8,677 focal audits conducted at 4,940 suppliers spanning 13 industries in 66 countries. The most common industries in our sample are garments, accessories, electronics, and toys (see Table 1). The majority of the audits took place in China; many of the rest took place elsewhere in Asia (Bangladesh, India, Indonesia, the Philippines, and Vietnam) and in North America (Mexico and the United States) (see Table 2). Auditors tend not to specialize by industry, but instead are assigned to audits largely based on their geographic proximity (to minimize travel costs and time) and their availability and to ensure that every audit team includes a trained lead auditor.[2]

[Insert Tables 1 and 2 here]

Dependent Variable

Our dependent variable measures a supplier's *improvement* between its prior and focal audits, which we calculate by dividing the number of violations from the focal audit plus 1 by the

[2] Nearly 80% of the auditors in our sample conducted audits in just one country. Of those who worked in more than one, most went only to nearby countries. The average auditor in our sample conducted audits of factories in nearly 5 of the 13 industries in our sample and nearly 25% conducted audits in 8 or more industries.

number of violations from the previous audit plus 1, then taking the natural log of that ratio, and then multiplying the result by – 1 so that higher values reflect greater improvement:

$$y_{i,t} = - 1 \times \ln [(V_{i,t} + 1) / (V_{i,t-1} + 1)],$$

where $V_{i,t}$ is the number of violations for supplier i audited at time t that pertain to child labor, forced or compulsory labor, working hours, occupational safety and health, minimum wage, treatment of foreign workers and subcontractors, and disciplinary practices and where $V_{i,t-1}$ is the comparable figure from the prior audit at time t-1.[3] We add 1 to both the numerator and the denominator to avoid losing observations in which either the current or prior audit yielded zero violations.[4] This metric, rather than the simple difference in violations, facilitates proportional comparisons between suppliers.[5] It also provides a more reliable estimate than a percent change metric, which can be overly sensitive to outliers and can inflate big changes.[6] Our log form reduces skewness[7] and enables a straightforward interpretation of our coefficients as elasticities. Multiplying the log ratio by – 1 results in larger values corresponding to greater improvement, which eases interpretation.

Independent Variables

The risk that the labor abuses documented in a social audit will be exposed and sanctioned depends on press freedom and NGO presence. We measure press freedom using the

[3] We excluded violations that, according to our data provider, do not apply to all suppliers (dormitory conditions, and canteen violations) or were interpreted differently by auditors in different countries (freedom of association, the right to organize and bargain collectively, legal or client requirements).
[4] Though only 4% of the prior audits in our sample had zero violations, such suppliers might be distinctively capable of exemplary performance and allowing these observations to drop out of the sample risks introducing bias.
[5] For example, our metric considers the proportional reduction from 12 to 6 violations at a large supplier to be equivalent to a small supplier's reduction from 4 to 2 violations, whereas a difference metric would consider the former to be three times the magnitude of the latter.
[6] For example, skewness declines by a factor of 10, from a value of 4.2 for percent change to a mere 0.4, for our *improvement* metric and kurtosis declines by a factor of nearly 7, from 30.5 for percent change to 4.5, for *improvement*.
[7] The simple ratio of violations at an establishment's focal audit to those at its prior audit is highly skewed: it ranges from 0 to 19, has a mean of 1.1, a standard deviation of 1.3, skewness of 4.2, and kurtosis of 30.5. Models that estimate this simple ratio as a dependent variable would be quite vulnerable to outliers driving their results.

Press Freedom Index from Reporters without Borders, which reflects the extent to which journalists in a given country faced direct and indirect threats such as imprisonment, physical attacks, and censorship, a metric used by others for the same purpose (e.g., Faccio, 2006; Cannizzaro and Weiner, 2015). We reverse-code the raw Press Freedom Index so that higher values indicate greater press freedom, rescale the result to range from 0 to 1, and take the log (after adding 1) to reduce skew. We measure NGO density as the number of NGOs in the supplier's country per million population—an approach used by others (e.g., Hafner-Burton and Tsutsui, 2005; Chih, Chih, and Chen, 2010; Jira and Toffel, 2013)—which we also log to reduce skew. We obtained NGO data from the Union of International Associations and population data from the US Census Bureau's International Data Base. Because both the press and NGOs are critical actors in the transnational advocacy networks we theorize will generate exposure risks, it is important to include both variables in our analysis. However, press freedom and NGO density are highly correlated (ρ=0.83). We therefore use principle components analysis as a data reduction technique (Hair et al., 1998: 95; Kennedy, 2008: 198), an approach others have used for the same purpose (e.g., Gulati and Sytch, 2007; Perkins, 2014; Guillén and Capron, 2015). The first component's eigenvalue of 1.85 is the only one to exceed the conventional threshold of 1 and it explains 92.3 percent of the variance between press freedom and NGO density. We refer to this first component as *exposure risk.*[8]

The magnitude of institutional pressure on suppliers depends not only on the probability of exposure, but also on their buyers' sensitivity to exposure. To operationalize that sensitivity, we rely on negative media reports, as others have done for similar purposes (Fiaschi, Giuliani, and Nieri, 2013; Wang and Ye, 2015; Kölbel, Busch, and Jancso, 2017). In particular, we

[8] A robustness test that includes *press freedom* and *NGO density* in our models instead of *exposure risk* yields the same inferences as our primary models.

consider whether a supplier serves a buyer that had recently been associated with supply chain labor abuses revealed in a news article or NGO report. To measure this, we relied on the database of media articles and reports on supply chain labor abuses compiled by the Business & Human Rights Resource Centre (BHRRC), an organization whose mission is to "amplify the voices of the vulnerable, and human rights advocates in civil society" by serving as a "global hub" of information collected by advocates in the field and by its own researchers (Business & Human Rights Resource Centre, 2017). BHRRC serves as a labor abuse information clearinghouse and research organization to "track the human rights policy and performance of over 7000 companies in over 180 countries, making information publicly available" (Business & Human Rights Resource Centre, 2017). It gathers news articles from around the world linking companies to human rights abuses and conveys this information to its 177,000 monthly website visitors and via its e-newsletter issued to thousands of activists, businesses, governments, global media, and investors who subscribe to its updates.[9] Because 80 percent of the observations in our sample have no such articles, we pursue a conservative approach to avoid outliers driving results by measuring *buyer exposure sensitivity (prior audit)* using a dummy (rather than a count) variable coded 1 for an audit of a supplier whose buyer was featured in at least one article in this database in the prior year, and coded 0 otherwise.

Whether an audit was expected or a surprise was measured by *announced*, a dichotomous variable coded 1 when the supplier had advance notice of the audit date and coded 0 for unannounced audits, based on data from the social auditor. Whether an audit is announced or

[9] The Business & Human Rights Resource Centre invites companies to respond to any post that names them and reports that 86% of companies do so (Page 2016), suggesting that companies are attentive to these reports.

unannounced is typically determined by the buyer on whose behalf the audit is being conducted. In our sample, 76 percent of the audits were announced.[10]

We calculate auditor training as the number of audit training courses an auditor had taken—which may address audit skills, specific audit topics, specific client issues, and certification schemes such as SA 8000—based on data provided by the social auditor.[11] Because audits are typically conducted by an audit team, we measure *maximum auditor training* as the largest number of training courses that any team member had undergone by the time the audit was conducted, which we log after adding 1 to reduce skew and then standardize to facilitate interpretation. The maximum number of training courses for audit teams averaged 6.9. We use the maximum training of any team member because this measures the greatest potential to identify code of conduct violations and to transfer knowledge on how to remediate them.[12]

Audit-level Control Variables

We control for audit-level factors by constructing variables from data provided by the social auditor. We control for the violations in the prior audit because suppliers whose prior audit yielded many violations face a different opportunity set than those with a "cleaner" history,

[10] We found no evidence that a supplier's prior violation count or the duration of a buyer-supplier relationship affected the propensity for a supplier's audit to be announced (versus unannounced). Specifically, we estimated a logistic regression that predicted whether an audit was announced (versus unannounced) based on the duration of the buyer-supplier relationship (proxied by whether an audit was the 2nd, 3rd, 4th, or 5th or more conducted of this supplier for the same buyer), buyer size (log employment) and country (dummies), controlling for supplier industry (dummies) and violations reported in the prior audit. Results indicate that buyer-supplier relationship is not a significant determinant of an audit being announced or unannounced. Larger buyers were more likely to have announced audits and the supplier-country dummies were jointly significant, as were the buyer-country dummies. The regression results that test our hypotheses (reported in Table 4) are unlikely to be contaminated by omitted variable bias associated with factors that predict whether an audit will be announced or unannounced because those regression models control for the statistically significant factors correlated with this decision.

[11] Training regarding audit skills helps auditors generally identify violations of codes of conduct. Training regarding specific audit topics includes courses on issues relevant to a specific industry, region, or supplier. Specific audit training could be about an issue such as child labor that could be a common problem for a particular supplier. Training on client issues educates auditors about certain codes of conduct the client is particularly concerned about or a specific protocol the client has agreed to comply with, such as SA8000.

[12] Robustness tests (not reported) indicate that using teams' average auditor training instead of maximum auditor training yields nearly identical results.

which may influence their likelihood of improvement. *Violations (prior audit)* is the number violations from a prior audit, top-coded at the 99th percentile of the sample distribution (25 violations) to reduce the potential impact of outliers and taking the log (after adding 1).

Because prior research indicates that auditing is less stringent when suppliers pay their own auditors (Jiang, Stanford, and Xie, 2012; Duflo et al., 2013; Short and Toffel, 2016), we created three dummy variables to indicate who paid for the audit: *paid by the supplier or third party, paid by the buyer* (on whose behalf the audit was conducted), and *paid by unknown entity* (to denote instances when we lacked information about who paid).

Re-audits typically have a more limited scope because they tend to focus on concerns raised at the prior audit. Because this could mechanically affect improvement rates, we include three dummy variables as controls: (1) *prior audit was re-audit, but focal audit was not*; (2) *focal audit was re-audit, but prior audit was not*; (3) *prior and focal audit were re-audits*. The baseline condition is that neither the prior nor the focal audit was a re-audit.[13]

Our interviews with social auditors—at the firm that provided us data and at others—indicated that the staff hours required to conduct an audit is a reasonable proxy for factory size and complexity, which could be associated with improvement but for which we lack direct measures. In addition, more staff hours in a prior audit might offer more opportunity to transfer information between the audit team and the supplier.[14] We therefore control for *audit duration*

[13] We have several reasons for believing that this approach adequately controls for the possible influence on improvement of either the prior or focal audit being a re-audit. We found no evidence that audits categorized as a re-audit averaged fewer violations than routine audits, whether based on a t-test, a simple negative binomial model, or a negative binomial model that also included industry dummies.

[14] We include in our model a dummy variable to denote the nearly 50% of observations for which the number of staff hours required to conduct the prior audit was missing from the database and where we thus recoded *auditor exposure (prior audit)* observations from missing values to 0. This common econometric approach is algebraically equivalent to recoding those missing values with the variable's mean (Greene, 2007: 62).

(prior audit), which we calculated by taking the log (after adding 1) of the number of staff hours required to conduct the prior audit, using data from the social auditing firm.

Audit teams including individuals who had previously audited the supplier have been shown to report fewer violations than teams whose members have no prior history there (Short, Toffel, and Hugill, 2016). We therefore created *previous auditor*, a dummy coded 1 when at least one member of the audit team had participated in the prior audit of that supplier and 0 otherwise. Because suppliers may remediate compliance problems identified at prior audits and thus face increasing mitigation costs, we create *audit sequence* as a counter variable to denote each supplier's first audit in the estimation sample, its second audit, and so on.[15] In our models, we flexibly control for *audit sequence* by including a dummy for each value, which avoids imposing the assumption that it has a linear influence on *improvement.*

Because an audit team's gender composition has been shown to affect audit results (Short, Toffel, and Hugill, 2016), we include three dummy variables: *all-female audit team*, *mixed-gender audit team*, and *all-male audit team.*[16]

We control for team experience, which has been shown to affect reported violations (Short, Toffel, and Hugill, 2016). We measure the *maximum auditor tenure* of each team as the maximum years of service with the social auditor among all team members. We include in our model both *maximum auditor tenure* and its squared value because the influence of experience on reported violations has been found to be nonlinear (Short, Toffel, and Hugill, 2016). As with

[15] A supplier's first observation in the estimation sample (where *audit sequence* equals 1) incorporates information from its focal audit (i,t) and prior audit (i,t-1) because our dependent variable incorporates both of their audit scores. *Audit sequence* equals 1 for 57% of the observations in our estimation sample, 2 for 22%, 3 for 10%, and 4 or more for 10%.

[16] 50% of our sample had all-female audit teams, 32% had all-male teams, and 15% had teams evenly split. The remaining 3% had teams that were 25%, 33%, 66%, or 75% female. Therefore, while we observe each team's precise gender composition, our dummies better reflect the distribution of our data than would a continuous percentage metric.

training, we measure this based on team members' maximum tenure to reflect the team's greatest potential to identify violations and to convey remediation ideas.[17]

Institutional Control Variables

Several factors pertaining to the supplier's institutional environment have been shown to affect violation rates (Toffel, Short, and Ouellet, 2015) and could affect improvement rates; we therefore control for them at the time of the prior audit.[18] A supplier country's dependence on foreign direct investment (FDI) might influence the extent to which the supplier perceives the need to respond to international pressure to improve how its factories are managed. We therefore control for each supplier country's percentage of gross domestic product (GDP) made up of FDI (*FDI inflows*) in the year of the prior audit, based on World Bank data.[19]

Because domestic legal protections for labor rights could influence how much pressure to improve suppliers perceive they are under, we obtain *labor laws* scores from Mosley (2011).[20] These scores measure the extent to which domestic law provides collective labor rights such as the rights to join unions and strike, whether government approval is required for collective bargaining, and whether laws restrict worker rights in export processing zones (Greenhill, Mosley, and Prakash, 2009). Because these scores are available only through 2002—before our sample period begins—we use the 2002 values for all years of our analysis. Studies have used

[17] Robustness tests (not reported) indicate that using teams' average tenure, rather than maximum tenure, yields nearly identical results.

[18] We only include the lagged value for these country-level variables because they are very stable over the period of time between two consecutive supplier audits. In particular, the correlation between *FDI inflows* at prior and focal audits is 0.90 and the correlation between *GDP per capita* at prior and focal audits is 0.99.

[19] *FDI inflows* measures net inflows of FDI (that is, inflows less divestment during the previous year) used to acquire a lasting management interest (that is, 10% or more of a company's voting stock was purchased by international entities) in the supplier's country. It is composed of equity capital, earnings reinvestment, and other short-term and long-term capital, as shown in the country's balance of payments.

[20] We find nearly identical results when, as a robustness test, we substitute for labor laws two alternative measures of the stringency of the domestic legal environment: the World Bank rule of law score and the number of ILO labor treaties the country has ratified.

this index to measure the stringency of country-level workers' rights protections generally (e.g., Greenhill, Mosley, and Prakash, 2009; Dean, 2015; Toffel et al., 2015; Fransen and Burgoon, 2017) on the basis that collective rights are "core" rights that are foundational to other workers' rights like wages, benefits, and working conditions (Greenhill, Mosley, and Prakash, 2009).

Because country-level wealth and differences in wealth between supplier and buyer countries could influence improvement rates, we control for *GDP per capita (prior audit)* and *GDP per capita in buyer country (prior audit)*, obtained from World Bank data. We control for potential differences in coercive pressure that buyers might exert based on their size by obtaining annual values of employment from Amadeus, Capital IQ, Hoovers, or Thomson ONE Banker, with which we create *log buyer employment (prior audit),* logging to reduce skew.[21]

Summary statistics are reported in Table 3.[22]

[Insert Table 3 here]

ESTIMATION AND RESULTS

We test our hypotheses by estimating a model that predicts *improvement* based on the independent and control variables described above and the following additional control variables. Whereas our hypothesized variables pertain to a supplier's prior audit, these same factors pertaining to the focal audit might influence the number of violations reported in that audit, which is used to construct our dependent variable. Since failing to account for these factors could bias our estimates, we also control for *announced (focal audit)* and *maximum auditor training (focal audit)*. Controlling for whether the focal audit is announced or unannounced prevents us from mistakenly attributing a supplier's reduction of focal-audit violations to situations in which

[21] The social auditor enabled us to append these variables to a list of buyer companies (which they provided to us without any other data) and subsequently provided the de-identified dataset.
[22] Correlations are reported in Table A1 in the Appendix.

advanced warning allowed it to fix or hide problems. Similarly, controlling for maximum auditor training at the focal audit prevents us from misattributing a supplier's reduction of violations to situations in which an establishment's focal audit team was less highly trained than its prior audit team. We do not include the focal-audit value of *exposure risk* because it is very stable over time—correlation between prior and focal audits is 0.99—and including it would substantially increase multicollinearity while adding little new information.

Because (a) several audit design elements and audit team characteristics at the prior audit could influence violations recorded in that audit and (b) these same factors at the focal audit could influence violations recorded in that audit, we include most audit-level controls—*paid by supplier or third party, paid by unknown entity, re-audit, previous auditor, all-female audit team, mixed-gender audit team,* and *maximum auditor tenure*—in the model twice to control for them at both the prior and focal audits.

We also include industry fixed effects and year fixed effects to control for potential differences in improvement rates between suppliers in different industries and between the years in our sample, respectively. Because suppliers might respond differently to buyers exerting varying levels of pressure based on the buyers' own institutional contexts (Toffel, Short, and Ouellet, 2015), we include fixed effects for buyer countries.[23] We log to reduce skew and then standardize *audit duration (prior audit)* and *maximum auditor training (prior audit)* to facilitate an elasticity interpretation of coefficients in response to a one-standard-deviation change and because we interact these variables with *announced*. We use the log form of all other continuous variables to facilitate their interpretations as elasticities.

[23] While we have 17 buyer countries in our sample, 89% of the observations correspond to just two. We therefore pursue a more conservative approach of including buyer-country fixed effects, controlling for differences in prosocial attitude in the buyer country as well as all other buyer country attributes that are relatively stable during our sample period.

Empirical Results

For context, we note that suppliers in our sample averaged 7.2 violations in their prior audit and 5.6 violations in their focal audit, an average improvement of 1.6 violations. This 22-percent improvement rate (calculated as 1.6 ÷ 7.2) corresponds to the sample average *improvement* rate of 0.22 reported in the summary statistics (Table 3).

We estimate our models using ordinary least squares (OLS) regression, clustering standard errors by the supplier's country, the most aggregated level of our explanatory variables.

[Insert Table 4 here]

We test Hypotheses 1–3 with Model 1, our baseline model, and report results in Table 4. The statistically significant positive coefficient on *exposure risk (prior audit)* ($\beta = 0.088$; $p < 0.01$) reveals that suppliers tend to improve more in countries with greater potential for civil society monitoring to expose noncompliance with codes of conduct, which supports H1. The coefficient magnitude indicates that a one-standard-deviation increase in *exposure risk (prior audit)* (such as a change from Vietnam to the Philippines) is associated with an increase in *improvement* from an average of 22 percent to 32.5 percent, based on average predictions across our sample.[24] This 32.5-percent improvement from the baseline average of 7.2 violations constitutes a reduction of 2.3 violations, which is nearly one and a half times the average reduction of 1.6.[25]

[24] 32.5% is calculated by adding to 0.22 the product of 0.088 (the coefficient on *exposure risk*) and 1.19 (the standard deviation of *exposure risk*).

[25] Our results are robust to replacing *risk exposure (prior audit)* with its two underlying elements: standardized *press freedom (prior audit)* and standardized *NGO density (prior audit)*. Akin to our primary results, this model's statistically significant negative coefficients on *press freedom* ($\beta = 0.373$; $p < 0.01$) and *NGO density (prior audit)* ($\beta = 0.036$; $p < 0.05$) indicate support for H1 that suppliers operating in institutional environments with greater press freedom and NGO pressure tend to improve more than suppliers in countries with less of those. The effect of *buyer exposure* is also unchanged, continuing to yield a positive statistically significant coefficient of a similar magnitude as in the primary model, supporting H2. Our results supporting H6a are similarly robust to the use of these alternative measures. Specifically, re-estimating Model 3 but interacting auditor training with either *press freedom* or *NGO density* instead of *exposure risk* yielded a coefficient on *announced (prior audit)* × *maximum auditor*

The statistically significant positive coefficient on *buyer exposure sensitivity (prior audit)* ($\beta = 0.067$; $p < 0.01$) illustrates greater average improvement for suppliers to buyers that have been publicly exposed for harms to workers in their supply chain, which supports H2. The coefficient magnitude indicates that suppliers audited on behalf of a buyer associated in the press with supply chain labor abuses experienced an average improvement of 27.5 percent, significantly more than the 20.8 percent improvement for other suppliers.

The statistically significant positive coefficient on *announced (prior audit)* in Model 1 ($\beta = 0.046$; $p < 0.01$) indicates that greater improvement follows announced audits than unannounced audits, which supports H3.[26] Predictive margins indicate that suppliers whose prior audit was announced experienced an average 23.1-percent improvement, compared to 18.5-percent for suppliers whose prior audit was unannounced. Applied to the average 7.23 violations in the prior audit, this translates to an average decline in 1.67 violations following announced audits versus 1.34 after unannounced audits. This average differential of 0.33 violations per audit corresponds to one additional violation being mitigated following three announced audits, compared to three unannounced audits.

The statistically significant positive coefficient on the standardized *maximum auditor training (prior audit)* ($\beta = 0.031$; $p < 0.01$) indicates that greater improvement tends to follow audits conducted by better-trained audit teams, which supports H4.[27] The coefficient magnitude

training (prior audit) that was significantly smaller than the coefficients on (a) *press freedom (prior audit)* × *maximum auditor training (prior audit)* (Wald $F = 9.37$; $p < 0.01$) and (b) *NGO density (prior audit)* × *maximum auditor training (prior audit)* (Wald $F = 17.49$; $p < 0.01$).

[26] Note that if announcing the prior audit gave factories time to hide or solve problems before auditors arrived, prior audits would yield fewer violations than they otherwise would, which would bias against our hypothesized result; the falsely depressed baseline violation count would make it more difficult to observe subsequent improvement.

[27] Our finding that better trained audit teams at the prior audit leads to more improvement would risk being driven by regression to the mean if our specification only measured audit team training associated with the prior audit. However, our models *also* control for the focal audit team's training, which mitigates that risk.

indicates that, on average, suppliers realize an additional 3.1-percentage-point improvement when their prior audit was conducted by a team whose best-trained auditor had one-standard-deviation more training than the average team's best-trained auditor (that is, 12.7 training courses versus the average of 6.9). Such suppliers average a 25.1-percent reduction (the sum of the 0.22 sample average and the 0.031 coefficient); a reduction of 1.8 violations from the prior to the focal audit, or 0.2 violations more than the average reduction of 1.6 violations.

To test Hypothesis 4, we add a term that interacts *announced (prior audit)* and *maximum auditor training (prior audit)* and report the results as Model 2 in Table 4. The statistically significant positive coefficient on the interaction term ($\beta = 0.049$; $p < 0.01$) indicates that better-trained audit teams at prior audits tend to prompt more improvement when those prior audits were announced than when they were unannounced, which supports H5. Figure 1 plots the average predicted effects of *maximum auditor training (prior audit)* on *improvement* for observations where prior audits were announced or unannounced. The upward-sloped dashed line indicates that for announced audits, better-trained auditors at the prior audit prompt more improvement. The relatively flat solid line indicates that for unannounced audits, suppliers' improvement rates are largely unaffected by how well-trained the prior audit team was.

Hypotheses 6a and 6b posit that auditor training will have a greater effect on improvement when audits are conducted in a cooperative manner, as indicated by an announced audit, than when audits threaten exposure and negative buyer reactions. Model 3 tests this by adding to the prior model an interaction term between *exposure risk (prior audit)* and *maximum auditor training (prior audit)*. A Wald test comparing the nonsignificant slightly negative coefficient on this interaction term ($\beta = -0.005$; $p = 0.42$) to the statistically significant positive coefficient on the interaction between *announced (prior audit)* and the standardized *maximum*

auditor training (prior audit) ($\beta = 0.048$; $p < 0.01$) indicates that these two coefficients significantly differ (Wald F = 22.08; $p < 0.01$). Model 4 tests H6b by interacting *maximum auditor training (prior audit)* with both *announced (prior audit)* and *buyer exposure sensitivity (prior audit)*. The results yield a nonsignificant slightly negative coefficient on the interaction between training and buyer exposure sensitivity ($\beta = -0.013$; $p = 0.22$) and a statistically significant positive coefficient on the interaction between *announced (prior audit)* and the standardized *maximum auditor training (prior audit)* ($\beta = 0.051$; $p < 0.01$). The latter is significantly larger than the former (Wald F = 13.39; $p < 0.01$), which supports H6b.

Supplementary Analysis

We conducted several additional analyses to assess the robustness of our results. While we believe our dependent variable is a well-designed interpretable metric robust to outliers, we acknowledge its complexity. Therefore, we assessed whether our results were sensitive to this metric by estimating models that instead predict the number of violations cited in the focal audit, controlling for the number of violations cited in the prior audit and including all other independent and control variables in our primary specifications. These negative binomial regression results, reported in Table A2 in the Appendix, confirm all inferences from our primary models and thereby indicate that our results are robust to this alternative specification.

Improvement might depend on the time between the prior and focal audits, especially if it were a particularly short or long time. To assess whether those more extreme cases influenced our results, we estimated our models on a subsample that excluded the 1,719 audits conducted fewer than 48 days or more than 433 days after the prior audit (the 10th and 90th percentiles of time span, respectively). The results were nearly identical to our main results. We also reestimated our models predicting an alternative dependent variable that explicitly accounted for

the time lag since the prior audit. Specifically, we predicted, *improvement rate*, the ratio of *improvement* (our primary dependent variable) to the log of the number days since the prior audit. The results continue to yield statistically significant support for all of our hypotheses.

We also estimated models with different sets of fixed effects. For example, we estimated models that included supplier-country fixed effects, omitting supplier-country variables to avoid multicollinearity because those variables exhibit little temporal variation (*exposure risk*, *FDI inflows*, *GDP per capita*) or no temporal variation (*labor laws (2002)*). The results yielded very similar coefficients and standard errors as our primary models and continue to support of all five audit-level hypotheses. We also estimated models that included buyer fixed effects to control for all time-invariant factors associated with buyers, omitting our buyer-level variables (*buyer exposure sensitivity* and *buyer employment*) to avoid multicollinearity. These models yielded coefficients and standard errors on our hypothesized variables very similar to those of our primary models, including on *exposure risk*, *auditor training*, and the interaction terms. The sole substantial difference is that the coefficient on *announced (prior audit)*, while still positive, is no longer statistically significant. This is at least partly due to the inclusion of buyer fixed effects cloaking much of the variation in *announced* and reveals that much of the statistical power driving the primary result (greater improvement after announced audit) is due to differences between suppliers rather than within suppliers.[28] We also estimated models that accounted for the fact that some buyers always sought unannounced audits, some always sought announced audits,

[28] 11% of the observations in our sample correspond to buyers who always called for unannounced audits and 54% to buyers that always called for announced audits. As a result, these fixed-effects models identify an effect of *announced (prior audit)* based only on the variation of the remaining 35% of observations, which pertain to buyers who sought a mix of announced and unannounced audits. Sample statistics indicate that improvement averages 0.056 among audits for buyers that always called for unannounced audits, but a much larger 0.257 among audits for buyers that always called for announced audits. That difference is far greater than the difference in average improvement among audits conducted for buyers that called for a mix of announced and unannounced audits, which we estimate as to be 0.209 following unannounced audits and 0.223 following announced audits.

and some sought a mix. Specifically, we added to our primary models one control variable that indicates whether an audit was conducted on behalf of a buyer that always specified announced inspections and another control variable that indicates buyers that always specified unannounced inspections (audits conducted on behalf of a buyer that specified a mix was the omitted category). The results are virtually identical to those of our primary models.

Our primary results are also robust to several sample restrictions. Some third-party social audit protocols detail specific procedures, such as whether audits should be announced or unannounced. To determine whether our results were driven by third-party protocol audits, we reestimate our models on a subsample excluding the 152 observations where the prior or focal audit was conducted according to a third-party protocol such as Business Social Compliance Initiative (BSCI), Customs-Trade Partnership Against Terrorism (C-TPAT), Social Accountability (SA) 8000, Sedex Members Ethical Trade Audit (SMETA), and Waste Reduction Audit Protocol (WRAP). The results were nearly identical to our main results.

Overall, our primary results proved markedly stable throughout various robustness tests.

Extensions

Our primary analysis measured improvement based on the difference between the total number of violations cited in the focal and prior audits, aggregating several types of violation to capture improvement comprehensively. Recognizing that different violation types may improve at different rates under the hypothesized conditions, we disaggregated our dependent variable by violation type to better understand which categories are more likely to improve under which conditions. To explore how our hypothesized variables influence these categories, we estimated models that predicted improvement in each violation category that made up our *improvement* variable for which at least 10 percent of audits exhibited variation: child labor, working hours,

minimum wage, and occupational safety and health (OSH).[29] We created an improvement metric for each category by applying the same formula used to create our primary improvement metric and we report in Table 5 the results of OLS regression models that predict each improvement metric based on the specifications used in our main models, except that our control for a supplier's prior violations corresponds to the specific violation category being predicted. The results broadly validate the mechanisms we theorize above, but indicate that our hypothesized variables are associated with varying degrees of improvement across different violation types and highlight important nuances underlying our primary results.

First, we find that under conditions of more exposure risk,[30] there is significantly more improvement in child labor, working hours, wages, and OSH violations. We also find more improvement in working hours and wages violations among suppliers of buyers with greater exposure sensitivity.[31] These are precisely the types of violation we would expect to improve if buyers are under institutional pressure to exhibit good supply chain practices. The more likely that highly embarrassing violations like child labor will be exposed, the more likely buyers are to pressure suppliers to correct them. Extensive overtime and insufficient worker payment often reflect demanding sourcing requirements that are out of suppliers' control. Buyers facing greater reputational risk from exposure may be less likely to make such demands.

---

[29] To avoid generalizing from their limited variation, we did not estimate models to predict the rarer changes (that is, in fewer than 10% of audits) in categories of *disciplinary practices, forced or compulsory labor* scores, and *treatment of foreign workers and subcontractors* scores.

[30] Average marginal effects indicate that a one-standard-deviation increase in *exposure risk (prior audit)* is associated with a 63% increase in child labor score improvement (calculated by dividing the coefficient, 0.044. by the dependent variable's mean, 0.07), a 68% increase in working hours improvement (0.034 ÷ 0.05), a 74% increase in minimum wage score improvement (0.059 ÷ 0.08), and a 17% increase in OSH score improvement (0.043 ÷ 0.25).

[31] Compared to suppliers of buyers that did not exhibit exposure sensitivity, suppliers of *buyers with exposure sensitivity (prior audit)* exhibited 79% more working hours score improvement (predictive margins indicate a 8.6% average working hours improvement rate versus a 4.8% average among suppliers whose prior audits were for buyers that did not exhibit exposure sensitivity) and 39% more minimum wage score improvement (10.6% versus 7.6%). We found no association between buyer exposure sensitivity and improvement in either child labor or OSH scores.

Second, we find that supplier improvement following announced audits is driven primarily by improvement in OSH violations.[32] These are precisely the kind of violation that may require transfer of knowledge between auditors and suppliers regarding buyer expectations and best practices. Notably, compliance with child labor restrictions improves significantly *less* following announced audits than unannounced audits. Because most factory managers are well aware that these are "zero-tolerance" violations that may be grounds for contract termination, they typically do not require information exchange to remedy; thus, we would not expect signals of greater trust to foster improvements in this area. In fact, our extension suggests that announcing audits could exacerbate such violations if it is taken as a signal of leniency rather than trust. This highlights the tradeoffs inherent in monitoring design choices, which may promote improvement in some categories but inhibit it in others.

Finally, we find that visits by highly trained auditors lead to significantly more improvement in child labor, wages, and OSH scores.[33] These findings are consistent with our hypothesis that better-trained auditors are better able to convey compliance information. Payroll and OSH practices can be complex and/or technical and a knowledgeable auditor can provide useful guidance about how to maintain effective record-keeping systems and remedy workplace hazards. Improvements in child labor scores may be attributable to suppliers' greater willingness to receive and follow advice from more highly trained auditors.

---

[32] Predictive margins indicate that suppliers' OSH scores improved 25.6% on average following announced audits, 18% more than the 21.6% average improvement following unannounced audits, and that child labor scores improved 6.9% following announced audits, 19% *less* than the average 8.2% improvement following unannounced audits. We found no association between a supplier's prior audit being announced (versus unannounced) and subsequent improvement in either its working hours or minimum wage scores.

[33] Average marginal effects indicate that a one-unit change in *maximum auditor training (prior audit)*—that is, a one-standard-deviation change in this standardized variable—is associated with an average 24% improvement in the child labor score (calculated by dividing the coefficient by the sample average dependent variable, or 0.017 ÷ 0.07), 43% improvement in the minimum wage score (0.034 ÷ 0.08), and 10% improvement in the OSH score (0.024÷ 0.25). We found no significant association between *auditor training (prior audit)* and improvement in working hour scores.

DISCUSSION

Taken together, our findings reveal structural conditions under which labor codes of conduct and monitoring regimes adopted by TNCs to manage the pressure of private political activism are, indeed, associated with improvements in their suppliers' labor practices. First, suppliers are more likely to improve when local and global institutional pressures generated by civil society activism create greater risk that harms to workers will be exposed. Civil society pressures in the supplier's local environment and reputational pressures on brands are both associated with greater improvement. These findings are consistent with literature on the importance of civil society actors in diffusing transnational norms (Keck and Sikkink, 1998) and suggest that private political activism targeting specific companies may be more effective when supported by ongoing activism that generates broad-based civil society pressures.

Second, we find suppliers improve more not only with external institutional pressures, but also with monitoring programs designed and implemented to encourage knowledge transfer. Suppliers improved more following audits with advance notice, particularly in areas like OSH, where compliance assistance can be most helpful. We also find suppliers more likely to improve when their auditors are highly trained, but only in the context of monitoring regimes in which audits are pre-announced. Highly trained auditors add no significant compliance improvement value through unannounced audits. Finally, while we find these two design features complementary, we find no evidence that institutional pressures amplify the positive association of auditor training with improvements in supplier labor practices.

In addition to theorizing and identifying conditions under which organizational structures adopted in response to private political activism will be associated with changes in organizational behavior, our study makes several further important contributions to the literature.

Private Politics and Coupling

This study answers calls to extend the literature on private politics from the realm of the symbolic into the realm of practical outcomes (De Bakker et al., 2013). Private politics scholarship tends to interpret corporate responses to activism as largely symbolic gestures (McDonnell and King, 2013), and it has done little to explore the conditions under which the organizational outcomes of private political activism could be "more than merely symbolic" (Short and Toffel, 2010). Indeed, the private politics literature has embraced the symbolism of corporate responses. Drawing on the social movements literature, it has been argued that symbolic outcomes are the appropriate focus of private politics research because changing market behavior is fundamentally a symbolic project of shaping political consciousness and constructing new understandings among market actors "about what is valuable and appropriate" (King and Pearce, 2010; McDonnell and King, 2013: 410). While this line of inquiry is important and illuminating, we argue that it should not inhibit investigation of the relationship between private political activism and actual changes in market behavior.

We undertake such an investigation by drawing on the decoupling literature and extending it in three ways designed to facilitate productive dialogue with the literature on private politics. First, we look beyond state-based institutions commonly found to promote coupling, demonstrating that organizational structures adopted in response to private political activism can improve organizational practices even where state regulation is weak. We identify factors associated with supplier compliance improvements in such contexts. Our focus on non-state structural contingencies yields insights that are particularly valuable to private political scholarship and practice, because institutional environments lacking robust state enforcement are the principal domain of private politics. Our finding that exposure risk created by NGOs and the

press is associated with tighter coupling of codes of conduct and supplier labor practices supports those who have argued that compliance pressure from civil society actors can be a supplement or substitute for government enforcement (e.g., Ayres and Braithwaite, 1992; Mattli and Woods, 2009). In addition, our finding of greater improvement among suppliers to reputation-compromised buyers suggests that exposing firms' corporate social irresponsibility can have enduring effects beyond those identified in the literature. Prior research has documented that global brands whose harmful practices are exposed by the media face risks including diminished brand value (Wang and Ye, 2015), credit risk (Kölbel, Busch, and Jancso, 2017), consumer boycotts, shareholder activism, more stringent government regulation, and lawsuits. Our finding that buyers tainted by negative media coverage become especially prone to working with suppliers that are more rapidly improving their working conditions suggests that these reputational risks may prompt firms to take substantive and not merely symbolic measures to avoid further reputation damage.

Second, our focus on structural contingencies associated with coupling extends qualitative research highlighting how micro-level individual action creates coupling contingencies. Furthermore, we address structural contingencies operating at two distinct levels—the institutional environment and program design—and we address their interaction. This multi-level approach extends the decoupling literature both empirically and theoretically. With few exceptions (e.g., Kalev, Dobbin, and Kelly, 2006), program design has been ignored as a coupling determinant and we are aware of no study that investigates how the efficacy of specific program design features depends on the institutional context in which they are implemented. We suspect that decoupling studies have devoted little time to the design features of organizational structures adopted in response to activism because this literature has long

theorized that such structures are likely to be merely symbolic. Our findings challenge that theoretical assumption by suggesting that the difference between symbol and substance may depend, in part, on how organizational structures are designed. Our interactions yield more nuanced insights into the conditions under which certain program design features are more likely to be associated with the coupling of program and practice. Cooperative signals enhanced the positive association between auditor training and improvement, while institutional pressures attenuated it, suggesting that knowledge transfer and exposure pressures may be alternative pathways to improvement. The larger point is that coupling contingencies are not necessarily additive; they can be interdependent and can amplify (or undermine) one another. Our findings suggest the crucial importance of understanding these interactions and the need to design monitoring regimes with the institutional context in mind.

These findings on the institutional and program design contingencies of coupling labor codes and practices are particularly illuminating for private politics. Lacking the state's coercive power, private political activists are limited in the corrective measures they can persuade companies to adopt. For instance, studies have found companies responding to activism by adopting measures like "impression management tactics" (McDonnell and King, 2013: 411), public "concessions" or commitments to conform to activists' demands (Eesley and Lenox, 2006; King, 2008), and organizational structures such as CSR officers or board committees (McDonnell, King, and Soule, 2015). The private politics literature tends to interpret these responses as largely symbolic, but our findings suggest that they need not inevitably be. Rather, it may be important to engage firms on the technical design elements of the structures they are being pushed to adopt. Program design and its intersection with politics and institutions points to an important new frontier for private political research and activism.

Third, our focus on *organizational change*—or supplier improvement—as our dependent variable highlights the importance of more precisely defining what coupling means and how that might change with empirical context. The decoupling literature has long been vague about what, empirically, constitutes coupling. Some identify coupling where an organizational structure is aligned with organizational practices at a single time (e.g., Westphal and Zajac, 2001; Zajac and Westphal, 2004; Lim and Tsutsui, 2012; Marquis and Qian, 2014), others where organizational structures are associated with changes in organizational practices over time (e.g., Cole, 2005, 2012; Short and Toffel, 2010). We suggest that neither approach is superior in all cases and that it is therefore important to match the empirical approach to the types of organizational structure studied and the purposes for which they are adopted. Because the point of private politics is to alter the behavior of market actors, it is essential to examine whether organizational structures adopted in responses to activism are associated with change.

Beyond the "Business Case" for Corporate Social Responsibility

Our study similarly extends existing research on CSR—the constellation of "organizational practices that are intended to serve stakeholders beyond the firm's owners, including employees, customers, communities, and society at large" (Chin, Hambrick, and Treviño, 2013: 202). A substantial body of research has sought to understand what drives companies to adopt CSR practices (e.g., Aguilera et al., 2007; David, Bloom, and Hillman, 2007; Marquis, Glynn, and Davis, 2007; Chiu and Sharfman, 2011; Chin, Hambrick, and Treviño, 2013; Briscoe, Gupta, and Anner, 2015) and how firms might benefit financially from CSR (e.g., Margolis and Walsh, 2003; Orlitzky, Schmidt, and Rynes, 2003; El Ghoul et al., 2011; Jiraporn et al., 2014). Our study moves beyond the "business case" for CSR to investigate whether the adoption of socially responsible practices like codes and monitoring is associated with

improvements in the *social* outcomes they purport to address—in this case, working conditions. Our work thus answers calls for more research into CSR's substantive impacts (Margolis and Walsh, 2003; Marquis and Qian, 2014; Briscoe, Gupta, and Anner, 2015).

Our finding that codes and monitoring are, under certain conditions, associated with improvements in supplier labor standards challenges the framing of normative debates about CSR. These debates have been highly polarized, with some espousing the view that CSR can address a variety of social harms and enhance societal well-being (e.g., Margolis and Walsh, 2003) while others dismiss it as a smokescreen to mask the insidious activities of TNCs (e.g., Banerjee, 2007). Our study calls for a more nuanced, contextual, and empirically grounded assessment of the virtues and vices of CSR.

Labor Standards in Global Supply Chains

Our findings also advance important, ongoing debates in the literature on supply chain labor standards regulation. First, we demonstrate that institutional pressures from civil society actors like NGOs and the press are associated with *improvements* in supplier labor practices and not merely indicative of the *level* of those practices at a particular point in time. Second, our finding that highly trained auditors are associated with greater improvement is a corrective to the literature's pervasive auditor-skepticism (O'Rourke, 2002; Esbenshade, 2004; Power, Ng, and Singh, 2008; Locke, Amengual, and Mangla, 2009; AFL-CIO, 2013; Karunakaran, 2013; LeBaron and Lister, 2015), suggesting the important role auditors can play given the proper tools.

Finally, our finding evidence that suppliers improved their OSH practices, but not their child labor practices, more following audits for which they had advance notice adds nuance to the debate surrounding whether audits should be announced or unannounced. The improvement

in OSH practices is consistent with economic models predicting that compliance can develop iteratively in response to trusting or cooperative gestures on the part of those implementing the rules (e.g., Axelrod, 1984; Scholz, 1984; Ayres and Braithwaite, 1992) and with qualitative studies finding better compliance with labor codes of conduct by suppliers in trusting and cooperative relationships with buyers (Frenkel and Scott, 2002; Locke and Romis, 2007). But our finding that announced audits impeded improvement in compliance with child labor standards suggests that cooperative signals from buyers may lead some suppliers to believe they can get away with such violations, consistent with research skeptical of the rigor of pre-announced audits (O'Rourke, 2002; Esbenshade, 2004; Clean Clothes Campaign, 2005; Gray, 2006; Power, Ng, and Singh, 2008; AFL-CIO, 2013; Karunakaran, 2013; LeBaron and Lister, 2015; Short, Toffel, and Hugill, 2016).

These findings highlight the difficult tradeoffs in designing monitoring regimes and the imperative that design choices reflect program aims. Monitoring programs should look different depending on whether they seek to collect the most complete and accurate information about supply chain conditions, to catch suppliers in the act of committing particularly harmful violations, or to improve working conditions through cooperative relationships with suppliers. In fact, some private regulatory programs have sought to reconcile these competing goals by strategically combining announced and unannounced audits. The ILO's Better Work program, for instance, undertakes some audits in an arms-length, policing fashion, but maintains a separate staff of monitors called "enterprise advisors" who engage in cooperative relations with factories to help improve working conditions. Also, our extension suggests that these tradeoffs could be reconciled to some degree through better training of auditors, having shown that highly trained auditors were associated with accelerated improvement in child labor practices as well as other

violations requiring compliance assistance. Further research should explore how different monitoring approaches can be deployed and combined to leverage their comparative advantages.

Compliance Knowledge Transfer in Monitoring Arrangements

Finally, our study highlights the relationship between knowledge transfer and regulatory compliance improvement and suggests fruitful avenues for research and theory bridging these domains. While the importance of compliance assistance and information exchange are often posited in research on supply chain standards (e.g., Locke, Amengual, and Mangla, 2009) and in the broader literature on regulatory governance (e.g., Sabel and Zeitlin, 2012; de Burca, Keohane, and Sabel, 2014), the role of information transfer in compliance improvement has not, to our knowledge, been empirically tested. Our findings that supplier compliance improves more following audits by highly trained auditors and that this training relationship is amplified in trust-signaling monitoring regimes strongly suggest that information transfer is a crucial mechanism for improving compliance. Future research is needed to document more directly the types of information exchanged between auditors and suppliers, the channels and methods of exchange, and their relationship to compliance improvements.

Such research will require more dialogue between the literatures on standards compliance and knowledge transfer. The latter has largely ignored the importance of information sharing in promoting regulatory compliance or compliance with social and environmental standards in supply chains. An extensive body of research on knowledge transfer examines the conditions under which information, expertise, and best practices can be transmitted between individuals, teams, and firms and within teams and firms (Hansen, 1999; Edmondson, Bohmer, and Pisano, 2001; Oxley and Wada, 2009), including supply chain partners (MacDuffie and Helper, 1997; Dyer and Chu, 2003; Kotabe, Martin, and Domoto, 2003; Modi and Mabert, 2007), but offers

little insight into how social auditors can best transfer compliance knowledge to audited firms. Indeed, this literature suggests many obstacles in that setting, including the geographic and cultural distance between buyers and suppliers (Szulanski, 1996; Almeida, Song, and Grant, 2002; Bhagat et al., 2002; Hernandez, 2014), the fact that the auditors who supply compliance information often have only fleeting social ties with the suppliers (Kotabe, Martin, and Domoto, 2003; Reagans and McEvily, 2003; Gomes-Casseres, Hagedoorn, and Jaffe, 2006), and the fact that the connection between code compliance and financial performance may not be immediately apparent to suppliers (who often fear it will erode their competitive advantage). Our study's suggestion that knowledge transfer can occur even under such conditions presents new opportunities for theorizing and empirical research.

Limitations and Future Research

Our study has limitations but also invites promising future research. First, because all the suppliers in our sample were audited, we address why some audited suppliers improve more rapidly than others, but not whether auditing is more effective than other possible interventions such as more stringent government regulation, legally binding international standards, or labor union activities. These are vital research questions.

Second, our findings are subject to several data limitations. Our study examines factories that faced at least two social audits by a single firm. By omitting establishments that were audited just once, we exclude audits that buyers might have initiated as a first step toward establishing a supplier relationship that was subsequently abandoned. Our focus on a single auditing firm has the advantage of providing comparable auditor training data, but does not enable us to compare practices across auditing firms. Omitting certain types of code violation from our analysis enhanced the reliability of our improvement measure, but leaves future

research to determine whether the factors we found to predict improvement would also do so with the types of violation we omitted, particularly those concerning freedom of association and collective bargaining. Data limitations also prevented us from controlling for some of the factors that prior studies have found to be predictive of regulatory compliance, such firm size and regulatory enforcement practices. While we believe that our proxies for key independent variables are reasonable, we cannot rule out the possibility that they are imperfect.

Conclusion

As private political strategies become increasingly important in pursuing social change, it is crucial to understand the full range of outcomes they can produce and under what conditions. We identify key structural conditions under which organizational structures adopted in response to private political activism measurably improve working conditions. Our findings suggest the need to look beyond the symbolism of organizational responses to assess how private political activism can improve corporate practices and social outcomes.

REFERENCES

Abito, J. M., D. Besanko, and D. Diermeier. 2016. "Corporate Reputational Dynamics, Private Regulation, and Activist Pressure." *Advances in Strategic Management* 34: 235–299.

AFL-CIO. 2013. *Responsibility Outsourced: Social Audits, Workplace Certification and Twenty Years of Failure to Protect*. http://www.aflcio.org/Learn-About-Unions/Global-Labor-Movement/Responsibility-Outsourced-Report (accessed December 18, 2014).

Aguilera, R. V., D. E. Rupp, C. A. Williams, and J. Ganapathi. 2007. "Putting the S Back in Corporate Social Responsibility: A Multilevel Theory of Social Change in Organizations." *Academy of Management Review* 32(3): 836–863.

Almeida, P., J. Song, and M. Grant. 2002. "Are Firms Superior to Alliances and Markets? An Empirical Test of Cross-Border Knowledge Building." *Organization Science* 13(2): 147–161.

Amengual, M. 2010. "Complementary Labor Regulation: The Uncoordinated Combination of State and Private Regulators in the Dominican Republic." *World Development* 38(3): 405–414.

Amengual M, and J. Chirot. 2016. "Reinforcing the State: Transnational and State Labor Regulation in Indonesia." *ILR Review* 69(5):1056–1080.

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.

Axelrod, R. 1984. *The Evolution of Cooperation*. New York: Basic.

Ayres, I., and J. Braithwaite. 1992. *Responsive Regulation: Transcending the Deregulation Debate*. New York: Oxford.

Banerjee, S. B. 2007. *Corporate Social Responsibility: The Good, the Bad, and the Ugly*. Northampton, MA: Edward Elgar.

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

Barkemeyer, R. 2009. "Beyond Compliance—Below Expectations? CSR in the Context of International Development." *Business Ethics: A European Review* 18(3): 273–289.

Baron, D. P. 2003. "Private Politics." *Journal of Economics & Management Strategy* 12(1): 31–66.

Barrientos, S., and S. Smith. 2007. "Do Workers Benefit from Ethical Trade? Assessing Codes of Labour Practice in Global Production Systems." *Third World Quarterly* 28(4): 713–729.

Bartley, T. 2005. "Corporate Accountability and the Privatization of Labor Standards: Struggles over Codes of Conduct in the Apparel Industry." *Research in Political Sociology* 14: 211–244.

Bartley, T. 2007. "How Foundations Shape Social Movements: The Construction of an Organizational Field and the Rise of Forest Certification." *Social Problems* 54(3): 229–255.

Bartley, T., and C. Child. 2014. "Shaming the Corporation: The Social Production of Targets and the Anti-Sweatshop Movement." *American Sociological Review* 79(4): 653–679.

Bartley, T., and N. Egels-Zandén. 2015. "Responsibility and Neglect in Global Production Networks: The Uneven Significance of Codes of Conduct in Indonesian Factories." *Global Networks* 15(S1): S21–S44.

Berliner, D., A. Regan Greenleaf, M. Lake, M. Levi, and J. Noveck. 2015. "Governing Global Supply Chains: What We Know (and Don't) about Improving Labor Rights and Working Conditions." *Annual Review of Law and Social Science* 11: 193–209.

Bhagat, R. S., B. L. Kedia, P. D. Harveston, and H. C. Triandis. 2002. "Cultural Variations in the Cross-Border Transfer of Organizational Knowledge: An Integrative Framework." *Academy of Management Review* 27(2): 204–221.

Boiral, O. 2007. "Corporate Greening Through ISO 14001: A Rational Myth?" *Organizational Science* 18(1):127–146.

Borgatti, S. P., and R. Cross. 2003. "A Relational View of Information Seeking and Learning in Social Networks." *Management Science* 49(4): 432–445.

Briscoe, F., A. Gupta, and M. S. Anner. 2015. "Social Activism and Practice Diffusion: How Activist Tactics Affect Non-targeted Organizations." *Administrative Science Quarterly* 60(2): 300–332.

Bromley P., and W. W. Powell. 2012. "From Smoke and Mirrors to Walking the Talk: Decoupling in the Contemporary World." *Academy of Management Annals* 6(1):483–530.

Business & Human Rights Resource Centre. 2017. "About Us." https://www.business-humanrights.org/en/about-us (accessed August 19, 2017).

Cannizzaro, A. P., and R. J. Weiner. 2015. "Multinational Investment and Voluntary Disclosure: Project-level Evidence from the Petroleum Industry." *Accounting, Organizations and Society* 42: 32–47.

Chatterji, A. K., and M. W. Toffel. 2010. "How Firms Respond to Being Rated." *Strategic Management Journal* 31(9): 917–945.

Cheng, J.-H., C.-H. Yeh, and C.-W. Tu. 2008. "Trust and Knowledge Sharing in Green Supply Chains." *Supply Chain Management: An International Journal* 13(4): 283–295.

Chih, H. L., H. H. Chih, and T. Y. Chen. 2010. "On the Determinants of Corporate Social Responsibility: International Evidence on the Financial Industry." *Journal of Business Ethics* 93(1): 115–135.

Chin, M. K., D. C. Hambrick, and L. K. Treviño. 2013. "Political Ideologies of CEOs: The Influence of Executives' Values on Corporate Social Responsibility." *Administrative Science Quarterly* 58(2): 197–232.

Chiu, S., and M. Sharfman. 2011. "Legitimacy, Visibility, and the Antecedents of Corporate Social Performance: An Investigation of the Instrumental Perspective." *Journal of Management* 37(6): 1558–1585.

Clean Clothes Campaign. 2005. "Looking for a Quick Fix: How Weak Social Auditing Is Keeping Workers in Sweatshops." https://www.evb.ch/fileadmin/files/documents/Konsum/Quickfix05.pdf (accessed May 31, 2016).

Cole, W. M. 2005. "Sovereignty Relinquished? Explaining Commitment to the International Human Rights Covenants, 1966–1999." *American Sociological Review* 70(3): 472–495.

Cole, W. M. 2012. "Human Rights as Myth and Ceremony? Reevaluating the Effectiveness of Human Rights Treaties, 1981–2007." *American Journal of Sociology* 117(4): 1131–1171.

Coleman, J. S. 1988. "Social Capital in the Creation of Human Capital." *American Journal of Sociology* 94(Supplement): S95–S120.

David, P., M. Bloom, and A. J. Hillman. 2007. "Investor Activism, Managerial Responsiveness, and Corporate Social Performance." *Strategic Management Journal* 28(1): 91–100.

Dean, Adam. 2015. "Power over Profits: The Political Economy of Workers and Wages." *Politics & Society* 43(3): 333–360.

De Bakker, F. G. A., F. den Hond, B. King, and K. Weber. 2013. "Social Movements, Civil Society and Corporations: Taking Stock and Looking Ahead." *Organization Studies* 34(5–6): 573–593.

de Burca, G., R. O. Keohane, and C. F. Sabel. 2014. "Global Experimentalist Governance." *British Journal of Political Science* 44(3): 477–486.

Deci, E. L., and R. M. Ryan. 1985. *Intrinsic Motivation and Self-Determination in Human Behavior*. New York: Plenum.

Dobbin, F., and E. L. Kelly. 2007. "How to Stop Harassment: Professional Construction of Legal Compliance in Organizations." *American Journal of Sociology* 112(4): 1203–1243.

Dobbin, F., and J. Sutton. 1998. "The Strength of a Weak State: The Employment Rights Revolution and the Rise of Human Resources Management Divisions." *American Journal of Sociology* 104(2): 441–476.

Dorobantu, S., W. J. Henisz, and L. Nartey. 2017. "Not All Sparks Light a Fire: Stakeholder and Shareholder Reactions to Critical Events in Contested Markets." *Administrative Science Quarterly* 62(3): 561–597.

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.

Dyer, J. H., and W. Chu. 2003. "The Role of Trustworthiness in Reducing Transaction Costs and Improving Performance: Empirical Evidence from the United States, Japan, and Korea." *Organization Science* 14(1): 57–68.

Edmondson, A. C., R. M. Bohmer, and G. P. Pisano. 2001. "Disrupted Routines: Team Learning and New Technology Implementation in Hospitals." *Administrative Science Quarterly* 46(4): 685–716.

Eesley, C., and M. J. Lenox. 2006. "Firm Responses to Secondary Stakeholder Action." *Strategic Management Journal* 27(8): 765–781.

Egels-Zandén, N. 2007. "Suppliers' Compliance with MNCs' Codes of Conduct: Behind the Scenes at Chinese Toy Suppliers." *Journal of Business Ethics* 75(1): 45–62.

Elevate. 2016. "ELEVATE." http://www.elevatelimited.com/ (accessed March 15, 2016).

El Ghoul, S., O. Guidhami, C. Kwok, and D. Mishra. 2011. "Does Corporate Social Responsibility Affect the Cost of Capital?" *Journal of Banking and Finance* 35(9): 2388–2406.

Esbenshade, J. 2004. *Monitoring Sweatshops: Workers, Consumers, and the Global Apparel Industry*. Philadelphia: Temple University Press.

Espeland, W. N. 1998. *The Struggle for Water: Politics, Rationality, and Identity in the American Southwest*. Chicago: University of Chicago.

Etzion, D., and F. Ferraro. 2010. "The Role of Analogy in the Institutionalization of Sustainability Reporting." *Organization Science* 21(5): 1092–1107.

Faccio, M. 2006. "Politically Connected Firms." *American Economic Review* 96(1): 369–386.

Fiaschi, D., E. Giuliani, and F. Nieri. 2013. "BRIC Companies in Search of Legitimacy: An Empirical Analysis of the Use of Different CSR Instruments." *Transnational Corporations* 22(3): 5–42.

Fransen, L. 2012. *Corporate Social Responsibility and Global Labor Standards: Firms and Activists in the Making of Private Regulation*. New York: Routledge.

Fransen, L., and B. Burgoon. 2017. "Support for Ethical Consumerism and Welfare States in the Global Economy: Complements or Substitutes?" *Global Policy* 8(S3): 42–55.

Frenkel, S. J., and D. Scott. 2002. "Compliance, Collaboration, and Codes of Labor Practice: The adidas Connection." *California Management Review* 45(1): 29–49.

Frynas, J. G. 2005. "The False Developmental Promise of Corporate Social Responsibility: Evidence from Multinational Oil Companies." *International Affairs* 81(3): 581–598.

Gereffi, G., J. Humphrey, and T. Sturgeon. 2005. "The Governance of Global Value Chains." *Review of International Political Economy* 12(1): 78–104.

Giambona, E., J. R. R. Graham, and C. R. R. Harvey. 2017. "The Management of Political Risk." *Journal of International Business Studies* 48(4): 523–533.

Gibbs, J. P. 1975. *Crime, Punishment, and Deterrence*. New York: Elsevier.

Gomes-Casseres, B., J. Hagedoorn, and A. B. Jaffe. 2006. "Do Alliances Promote Knowledge Flows?" *Journal of Financial Economics* 80(1): 5–33.

Gould, D. 2005. "The Problem with Supplier Audits." *Corporate Responsibility Management* 2(1): 24–29.

Gray, G. C. 2006. "The Regulation of Corporate Violations: Punishment, Compliance, and the Blurring of Responsibility." *British Journal of Criminology* 46(5): 875–892.

Greene, W. H. 2007. *Econometric Analysis,* 6th Edition. Upper Saddle River, NJ: Prentice Hall.

Greenhill, B., L. Mosley, and A. Prakash. 2009. "Trade-based Diffusion of Labor Rights: A Panel Study, 1986–2002." *American Political Science Review* 103(4): 169–190.

Guillén, M. F., and L. Capron. 2015. "State Capacity, Minority Shareholder Protections, and Stock Market Development." *Administrative Science Quarterly* 61(1): 125–160.

Gulati, R., and M. Sytch. 2007. "Dependence Asymmetry and Joint Dependence in Interorganizational Relationships: Effects of Embeddedness on a Manufacturer's Performance in Procurement Relationships." *Administrative Science Quarterly* 52(1): 32–69.

Hafner-Burton, E. M., and K. Tsutsui. 2005. "Human Rights in a Globalizing World: The Paradox of Empty Promises." *American Journal of Sociology* 110(5):1373–1411.

Hair, J. F., Jr., R. E. Anderson, R. L. Tatham, and W. C. Black. 1998. *Multivariate Data Analysis,* 5th Edition. Delhi, India: Pearson Education.

Hallett, T. 2010. "The Myth Incarnate: Recoupling Processes, Turmoil, and Inhabited Institutions in an Urban Elementary School." *American Sociological Review* 75(1): 52–74.

Hansen, M. T. 1999. "The Search-Transfer Problem: The Role of Weak Ties in Sharing Knowledge across Organization Subunits." *Administrative Science Quarterly* 44(1): 82–111.

Hernandez, E. 2014. "Finding a Home Away from Home: Effect of Immigrants on Firms' Foreign Location Choice and Performance." *Administrative Science Quarterly* 59(1): 73–108.

Inkpen, A. C., and E. W. K. Tsang. 2005. "Social Capital, Networks, and Knowledge Transfer." *Academy of Management Review* 30(1): 146–165.

Intertek. 2016. “Global Supplier Management System.” http://www.intertek.com/consumer/auditing/global-supplier-management/ (accessed May 31, 2016).

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.

Jira, C., and M. W. Toffel. 2013. “Engaging Supply Chains in Climate Change.” *Manufacturing and Service Operations Management* 15(4): 559–577.

Jiraporn, P., N. Jiraporn, A. Boeprasert, and K. Chang. 2014. “Does Corporate Social Responsibility (CSR) Improve Credit Ratings? Evidence from Geographic Identification.” *Financial Management* 43(3): 505–531.

Justice, D. W. 2006. “Corporate Social Responsibility and Society’s Expectations of Business.” *Perspectives on Work* 9(2): 14–16.

Kalev, A., F. Dobbin, and E. Kelly. 2006. “Best Practices or Best Guesses? Assessing the Efficacy of Corporate Affirmative Action and Diversity Policies.” *American Sociological Review* 71(4): 589–617.

Karunakaran, N. 2013. “Retailers Like Marks & Spencer, Walmart Readies to Address Human-Rights Issues of Workers.” *Economic Times* (December 11). http://articles.economictimes.indiatimes.com/2013-12-10/news/45035505_1_marks-spencer-supply-chains-factories (accessed May 4, 2016).

Keck, M. E., and K. Sikkink. 1998. *Activists beyond Borders*. Ithaca, NY: Cornell.

Kennedy, P. 2008. *A Guide to Econometrics,* 6th Edition*.* Malden, MA: Wiley-Blackwell.

King, B. G. 2008. “A Political Mediation Model of Corporate Response to Social Movement Activism.” *Administrative Science Quarterly* 53(3): 395–421.

King, B. G. 2014. “Reputational Dynamics of Private Regulation.” *Socio-Economic Review* 12(1): 200–206.

King, B. G., and N. A. Pearce. 2010. “The Contentiousness of Markets: Politics, Social Movements, and Institutional Change in Markets.” *Annual Review of Sociology* 36: 249–267.

King, B. G., 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.

Kölbel, J. F., T. Busch, and L. M. Jancso. 2017. “How Media Coverage of Corporate Social Irresponsibility Increases Financial Risk.” *Strategic Management Journal* 38 (11): 2266–2284.

Kotabe, M., X. Martin, and H. Domoto. 2003. “Gaining from Vertical Partnerships: Knowledge Transfer, Relationship Duration, and Supplier Performance Improvement in the U.S. and Japanese Automotive Industries.” *Strategic Management Journal* 24(4): 293–316.

Kotchen, M., and J. J. Moon. 2012. “Corporate Social Responsibility for Irresponsibility” *B.E. Journal of Economic Analysis & Policy (Contributions)* 12(1), Article 55.

LeBaron, G., and J. Lister. 2015. “Benchmarking Global Supply Chains: The Power of the ‘Ethical Audit’ Regime.” *Review of International Studies* 41(Special issue 5): 905–924.

Levine, D. I., M. W. Toffel, and M. S. Johnson. 2012. “Randomized Government Safety Inspections Reduce Worker Injuries with No Detectable Job Loss.” *Science* 336(6083): 907–911.

Lim, A., and K. Tsutsui. 2012. “Globalization and Commitment in Corporate Social Responsibility: Cross-national Analyses of Institutional and Political-economy Effects.” *American Sociological Review* 77(1): 69–98.

Locke, R., M. Amengual, and A. Mangla. 2009. “Virtue Out of Necessity? Compliance, Commitment, and the Improvement of Labor Conditions in Global Supply Chains.” *Politics & Society* 37(3): 319–351.

Locke, R. M. 2013. *The Promise and Limits of Private Power: Promoting Labor Standards in a Global Economy*. Cambridge, UK: Cambridge University Press.

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. Rissing, and T. Pal. 2013. "Complements or Substitutes? Private Codes, State Regulation and the Improvement of Labor Standards in Global Supply Chains." *British Journal of Industrial Relations* 51(3): 519–552.

Locke, R. M., and M. Romis. 2007. "Improving Work Conditions in a Global Supply Chain." *MIT Sloan Management Review* 48(2): 54–62.

Loewenstein, J., L. Thompson, and D. Gentner. 2003. "Analogical Learning in Negotiation Teams: Comparing Cases Promotes Learning and Transfer." *Academy of Management Learning and Education* 2(2): 119–127.

Luo, J., S. Meier, and F. Oberholzer-Gee. 2012. "No News Is Good News: CSR Strategy and Newspaper Coverage of Negative Firm Events." Harvard Business School Working Paper 12-091.

MacDuffie, J. P., and S. Helper. 1997. "Creating Lean Suppliers: Diffusing Lean Production through the Supply Chain." *California Management Review* 39(4): 118–151.

Margolis, J. D., and J. P. Walsh. 2003. "Misery Loves Companies: Rethinking Social Initiatives by Business." *Administrative Science Quarterly* 48(2): 268–305.

Marks, K. 2012. "Factory Workers Are 'Forced to Lie' During Adidas Safety Inspections." *Independent (UK)* (April 13). http://www.independent.co.uk/news/world/asia/factory-workers-are-forced-to-lie-during-adidas-safety-inspections-7644018.html (accessed May 4, 2016).

Marquis, C., M. A. Glynn, and G. F. Davis. 2007. "Community Isomorphism and Corporate Social Action." *Academy of Management Review* 32(3): 925–945.

Marquis, C., and C. Qian. 2014. "Corporate Social Responsibility Reporting in China: Symbol or Substance?" *Organization Science* 25(1): 127–148.

Mattli, W., and N. Woods. 2009. "In Whose Benefit? Explaining Regulatory Change in Global Politics." In W. Mattli and N. Woods (eds.), *The Politics of Global Regulation*: 1–43. Princeton, NJ: Princeton University Press.

Mayer, R. C., J. H. Davis, and F. D. Schoorman. 1995. "An Integrative Model of Organizational Trust." *Academy of Management Review* 20(3): 709–734.

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.

McDonnell, M.-H., and B. G. King. 2013. "Keeping Up Appearances: Reputation Threat and Prosocial Responses to Social Movement Boycotts." *Administrative Science Quarterly* 58(3): 387–419.

McDonnell, M.-H., B. G. King, and S. A. Soule. 2015. "A Dynamic Process Model of Private Politics: Activist Targeting and Corporate Receptivity to Social Challenges." *American Sociological Review* 80(3): 654–678.

McEvily, B., V. Perrone, and A. Zaheer. 2003. "Trust as an Organizing Principle." *Organization Science* 14(1): 91–103.

Mena, S., and D. Waeger. 2014. "Activism for Corporate Responsibility: Conceptualizing Private Regulation Opportunity Structures." *Journal of Management Studies* 51(7): 1091–1117.

Meyer, J. W., J. Boli, G. M. Thomas, and F. O. Ramirez. 1997. "World Society and the Nation-State." *American Journal of Sociology* 103(1):144–181.

Meyer, J. W., and B. Rowan. 1977. "Institutionalized Organizations: Formal Structure as Myth and Ceremony." *American Journal of Sociology* 83(2): 340–363.

Modi, S. B., and V. A. Mabert. 2007. "Supplier Development: Improving Supplier Performance through Knowledge Transfer." *Journal of Operations Management* 25(1): 42–64.

Mosley, L. 2011. "Replication Data for: Collective Labor Rights Dataset." http://hdl.handle.net/1902.1/15590 (accessed August 29, 2013).

Nadler, J., L. Thompson, and L. Van Boven. 2003. "Learning Negotiation Skills: Four Models of Knowledge Creation and Transfer." *Management Science* 49(4): 529–540.

Nadvi, K., P. Lund-Thomsen, H. Xue, and N. Khara. 2011. "Playing against China: Global Value Chains and Labour Standards in the International Sports Goods Industry." *Global Networks* 11(3): 334–354.

Obstfeld, D. 2005. "Social Networks, the *Tertius Iungens* Orientation, and Involvement in Innovation." *Administrative Science Quarterly* 50(1): 100–130.

Oka, C. 2010a. "Accounting for the Gaps in Labour Standard Compliance: The Role of Reputation-conscious Buyers in the Cambodian Garment Industry." *European Journal of Development Research* 22(1): 59–78.

Oka, C. 2010b. "Channels of Buyer Influence and Labor Standard Compliance: The Case of Cambodia's Garment Sector." In D. Lewin, B. E. Kaufman, and P. J. Gollan (eds.), *Advances in Industrial and Labor Relations* 17: 153–183. Bingley, UK: Emerald Group.

Orlitzky, M., F. L. Schmidt, and S. L. Rynes. 2003. "Corporate Social and Financial Performance: A Meta-analysis." *Organization Studies* 24(3): 403–441.

O'Rourke, D. 2002. "Monitoring the Monitors: A Critique of Corporate Third-party Labour Monitoring." In R. Jenkins, R. Pearson, and G. Seyfang (eds.), *Corporate Responsibility and Labour Rights: Codes of Conduct in the Global Economy*: 196–207. London: Earthscan.

Overdevest, C. 2010. "Comparing Forest Certification Schemes: The Case of Ratcheting Standards in the Forest Sector." *Socio-Economic Review* 8(1): 47–76.

Oxley, J., and T. Wada. 2009. "Alliance Structure and the Scope of Knowledge Transfer: Evidence from U.S.-Japan Agreements." *Management Science* 55(4): 635–649.

Page, A. M. 2016. "The Alchemy of Business and Human Rights (Part IV): The Rules of the Game." *Huffington Post* (May 11). https://www.huffingtonpost.com/aaron-marr-page/the-alchemy-of-business-h_2_b_9852888.html (accessed May 16, 2018).

Perkins, S. E. 2014. "When Does Prior Experience Pay? Institutional Experience and the Multinational Corporation." *Administrative Science Quarterly* 59(1): 145–181.

Plambeck, E. L., and T. A. Taylor. 2016. "Supplier Evasion of a Buyer's Audit: Implications for Motivating Supplier Social and Environmental Responsibility." *Manufacturing & Service Operations Management* 18(2): 184–197.

Potoski, M., and A. Prakash. 2005. "Green Clubs and Voluntary Governance: ISO 14001 and Firms' Regulatory Compliance." *American Journal of Political Science* 49(2): 235–248.

Power, C., I.-C. Ng, and M. Singh. 2008. "The Burden of Good Intentions." *Time International* 171(25): 45–48.

Reagans, R., and B. McEvily. 2003. "Network Structure and Knowledge Transfer: The Effects of Cohesion and Range." *Administrative Science Quarterly* 48(2): 240–267.

Reinholt, M., T. Pederson, and N. J. Foss. 2011. "Why a Central Network Position Isn't Enough: The Role of Motivation and Ability for Knowledge Sharing in Employee Networks." *Academy of Management Journal* 54(6): 1277–1297.

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.* Geneva: United Nations, Human Rights Council. http://www.reports-and-materials.org/Ruggie-report-7-Apr-2008.pdf (accessed July 5, 2015).

Ruwanpura, K. N. 2012. "Ethical Codes: Reality and Rhetoric. A Study of Sri Lanka's Apparel Sector" University of Southampton working paper. http://eprints.soton.ac.uk/337113/ (accessed May 31, 2016).

Sabel, C. F., and J. Zeitlin. 2012. "Experimentalist Governance." In D. Levi-Faur (ed.), *The Oxford Handbook of Governance*: 169–183. Oxford, UK: Oxford University Press.

Scholz, J. T. 1984. "Cooperation, Deterrence, and the Ecology of Regulatory Enforcement." *Law & Society Review* 18(2): 179–224.

Seidman, G. 2007. *Beyond the Boycott: Labor Rights, Human Rights, and Transnational Activism.* New York: Russell Sage Foundation.

Shamir, R. 2011. "Socially Responsible Private Regulation: World-Culture or World-Capitalism?" *Law & Society Review* 45(2): 313–336.

Shea, A., M. Nakayama, and J. Heymann. 2010. "Improving Labour Standards in Clothing Factories: Lessons from Stakeholder Views and Monitoring Results in Cambodia." *Global Social Policy* 10(1): 85–110.

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.

Short, J. L., and M. W. Toffel. 2016. "The Integrity of Private Third-party Compliance Monitoring." *Administrative & Regulatory Law News* 42(1): 22–25.

Short, J. L., M. W. Toffel, and A. R. Hugill. 2016. "Monitoring Global Supply Chains." *Strategic Management Journal* 37(9): 1878–1897.

Suchman, Mark C. 1995. "Managing Legitimacy: Strategic and Institutional Approaches." *Academy of Management Review* 20(3): 571–610.

Szulanski, G. 1996. "Exploring Internal Stickiness: Impediments to the Transfer of Best Practice within the Firm." *Strategic Management Journal* 17(S2): 27–43.

Tenbrunsel, A. E., and D. M. Messick. 1999. "Sanctioning Systems, Decision Frames, and Cooperation." *Administrative Science Quarterly* 44(4): 684–707.

Thomas-Hunt, M. C., T. Y. Ogden, and M. A. Neale. 2003. "Who's Really Sharing? Effects of Social and Expert Status on Knowledge Exchange within Groups." *Management Science* 49(4): 464–477.

Thompson, L., D. Gentner, and J. Loewenstein. 2000. "Avoiding Missed Opportunities in Managerial Life: Analogical Training More Powerful Than Individual Case Training." *Organization Behavior and Human Decision Processes* 82(1): 60–75.

Tilcsik, A. 2010. "From Ritual to Reality: Demography, Ideology, and Decoupling in a Post-Communist Government Agency." *Academy of Management Journal* 53(6): 1474–1498.

Toffel, M. W., J. L. Short, and M. Ouellet. 2015. "Codes in Context: How States, Markets, and Civil Society Shape Adherence to Global Labor Standards." *Regulation & Governance* 9(3): 205–223.

UL Responsible Sourcing. 2015. "Supply Chain Certification." http://services.ul.com/service/supply-chain-certification/ (accessed May 15, 2016).

Utting, P. 2005. "Corporate Responsibility and the Movement of Business." *Development in Practice* 15(3–4): 375–388.

Wang, J., and K. Ye. 2015. "Media Coverage and Firm Valuation: Evidence from China." *Journal of Business Ethics* 127(3): 501–511.

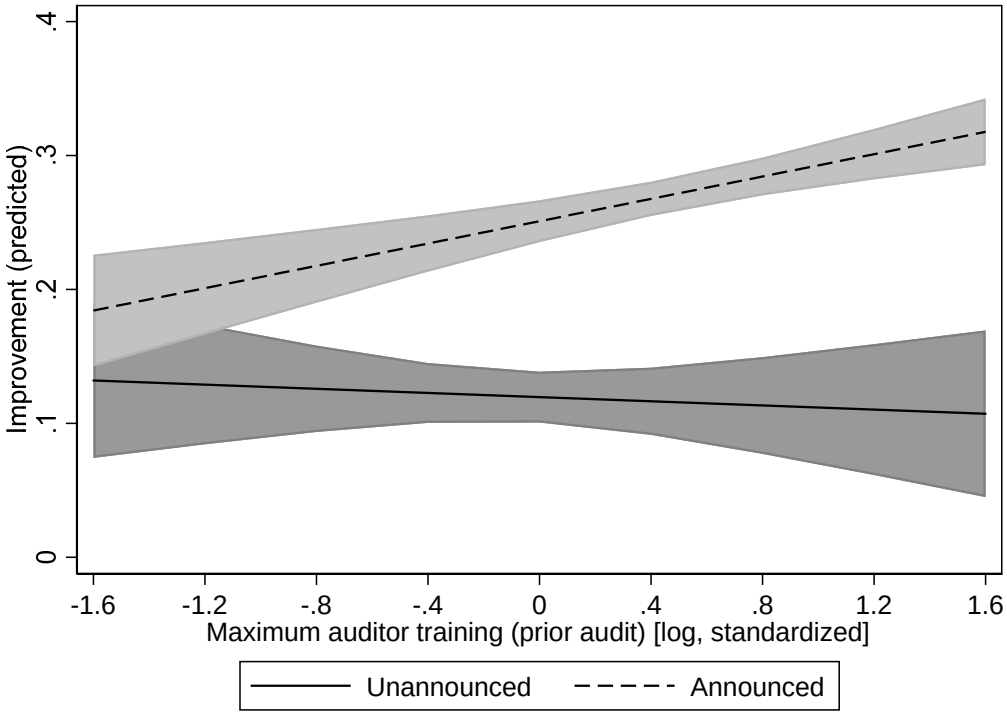
Weil, D., and C. Mallo. 2007. "Regulating Labour Standards via Supply Chains: Combining Public/Private Interventions to Improve Workplace Compliance." *British Journal of Industrial Relations* 45(4): 791–814.

Westphal, J. D., and E. J. Zajac. 2001. "Decoupling Policy from Practice: The Case of Stock Repurchase Programs." *Administrative Science Quarterly* 46(2): 202–228.

Zajac, E. J., and J. D. Westphal. 2004. "The Social Construction of Market Value: Institutionalization and Learning Perspectives on Stock Market Reactions." *American Sociological Review* 69(3): 433–457.

Zajak, S. *Transnational Activism, Global Labor Governance, and China*. New York: Palgrave Macmillan.

Figure 1. Average Predicted Improvement Values Based on Varying Amounts of Maximum Auditor Training at Prior Audits That Were Unannounced or Announced



Note: Average predicted values of *improvement* from Model 2 of Table 4 when the prior audit was announced (solid line) or unannounced (dashed line), based on varying levels of *maximum auditor training (prior audit)* (logged and standardized) and all other variables at their actual values.

Table 1. Industry Composition of Audits and Audited Suppliers

	Audits		Suppliers	
	Number	Percent	Number	Percent
Accessories	930	10.7%	506	10.2%
Building materials	143	1.6%	74	1.5%
Chemicals and plastics	47	0.5%	36	0.7%
Electronics	358	4.1%	171	3.5%
Food, agriculture, beverages	73	0.8%	56	1.1%
Footwear	191	2.2%	103	2.1%
Furniture	226	2.6%	103	2.1%
Garment	2,902	33.4%	1,659	33.6%
Metal products	85	1.0%	45	0.9%
Paper, printing, and publishing	117	1.3%	71	1.4%
Services	25	0.3%	15	0.3%
Toys	269	3.1%	139	2.8%
Unknown (other and missing)	3,311	38.2%	1,962	39.7%
Total	8,677	100.0%	4,940	100.0%

Table 2. Location of Audits and Audited Suppliers

	Audits		Suppliers	
Supplier country	Number	Percent	Number	Percent
Bangladesh	129	1.5%	104	2.1%
Brazil	40	0.5%	32	0.6%
Canada	33	0.4%	19	0.4%
China (includes Hong Kong)	6,345	73.1%	3,416	69.1%
Egypt	23	0.3%	13	0.3%
Guatemala	29	0.3%	24	0.5%
India	288	3.3%	207	4.2%
Indonesia	143	1.6%	101	2.0%
Italy	30	0.3%	28	0.6%
Jordan	36	0.4%	22	0.4%
Korea, Republic of (South)	53	0.6%	39	0.8%
Malaysia	40	0.5%	25	0.5%
Mexico	77	0.9%	66	1.3%
Pakistan	75	0.9%	53	1.1%
Peru	35	0.4%	23	0.5%
Philippines	121	1.4%	71	1.4%
Sri Lanka	55	0.6%	40	0.8%
Thailand	41	0.5%	30	0.6%
Turkey	84	1.0%	59	1.2%
United States	629	7.2%	302	6.1%
Vietnam	187	2.2%	121	2.4%
Countries with <20 audits in sample	184	2.1%	145	2.9%
Total	8,677	100.0%	4,940	100.0%

Table 3. Summary Statistics

	Mean	SD	Min	Max
Improvement	0.22	0.86	-3.26	3.26
Exposure risk (prior audit)	-0.34	1.19	-1.14	4.02
Press freedom§ (prior audit)	0.00	1	-0.77	2.37
NGO density§ (prior audit)	0.00	1	-0.55	4.85
Buyer exposure sensitivity (prior audit)	0.20	0.40	0	1
Announced (prior audit)	0.76	0.43	0	1
Maximum auditor training§ (prior audit)	0.00	1	-1.65	1.64
Violations (prior audit)	7.23	5.81	0	25
Violations† (prior audit)	1.84	0.77	0	3.26
Announced (focal audit)	0.77	0.42	0	1
Maximum auditor training† (focal audit)	2.04	0.87	0	3.47
Paid by supplier or third party (prior audit)	0.51	0.50	0	1
Paid by the buyer (prior audit)	0.43	0.49	0	1
Paid by unknown entity (prior audit)	0.06	0.24	0	1
Paid by supplier or third party (focal audit)	0.52	0.50	0	1
Paid by the buyer (focal audit)	0.42	0.49	0	1
Paid by unknown entity (focal audit)	0.05	0.22	0	1
Prior audit was re-audit, but focal audit was not	0.10	0.30	0	1
Focal audit was re-audit, but prior audit was not	0.38	0.49	0	1
Prior and focal audit were re-audits	0.19	0.40	0	1
Audit duration† (prior audit)	1.44	0.43	0.03	4.39
Previous auditor (prior audit)	0.20	0.40	0	1
Previous auditor (focal audit)	0.23	0.42	0	1
Audit sequence	1.84	1.31	1	11
All-female audit team (prior audit)	0.53	0.50	0	1
Mixed-gender audit team (prior audit)	0.16	0.37	0	1
All-female audit team (focal audit)	0.50	0.50	0	1
Mixed-gender audit team (focal audit)	0.18	0.38	0	1
Maximum auditor tenure† (prior audit)	1.86	0.28	0.69	2.77
Maximum auditor tenure† (focal audit)	1.78	0.32	0.69	2.77
Buyer employment † (prior year)	10.32	1.22	1.61	14.35
FDI inflows† (prior audit)	1.54	0.38	-0.45	3.41
Labor laws† (2002)	3.12	0.12	2.30	3.35
GDP per capita† (prior audit)	7.78	1.00	5.61	10.68
GDP per capita in buyer country† (prior audit)	10.59	0.26	6.59	10.85

Notes: † indicates logged. § indicates logged, then standardized. N=8,677 audits, except 7,774 for *buyer exposure sensitivity*, 4,338 for *audit duration (prior audit)*, 4,523 for *previous auditor (prior audit)*, 8,668 for *previous auditor (focal audit)*, and 5,355 for *buyer employment*.

Table 4. Regression Results

Dependent variable: Improvement

	(1)	(2)	(3)	(4)
Exposure risk (prior audit)	0.088**	0.088**	0.086**	0.088**
	[0.015]	[0.015]	[0.016]	[0.015]
Buyer exposure sensitivity (prior audit)	0.067**	0.071**	0.071**	0.073**
	[0.021]	[0.020]	[0.021]	[0.020]
Announced (prior audit)	0.046**	0.046**	0.046**	0.045**
	[0.011]	[0.013]	[0.013]	[0.013]
Maximum auditor training§ (prior audit)	0.031**	-0.008	-0.009	-0.002
	[0.011]	[0.018]	[0.016]	[0.017]
Maximum auditor training§ (prior audit) × Announced (prior audit)		0.049**	0.048**	0.051**
		[0.012]	[0.012]	[0.013]
Maximum auditor training§ (prior audit) × Exposure risk (prior audit)			-0.005	
			[0.006]	
Maximum auditor training§ (prior audit) × Buyer exposure sensitivity (prior audit)				-0.013
				[0.010]
Announced (focal audit)	0.049**	0.048*	0.047*	0.046*
	[0.018]	[0.018]	[0.018]	[0.018]
Maximum auditor training† (focal audit)	-0.052**	-0.053**	-0.053**	-0.053**
	[0.016]	[0.016]	[0.016]	[0.016]
Violations† (prior audit)	0.675**	0.678**	0.679**	0.679**
	[0.026]	[0.027]	[0.027]	[0.027]
Paid by supplier or third party (prior audit)	-0.018	-0.022	-0.021	-0.023
	[0.021]	[0.020]	[0.021]	[0.021]
Paid by supplier or third party (focal audit)	0.139**	0.144**	0.144**	0.146**
	[0.018]	[0.018]	[0.018]	[0.018]
Paid by unknown entity (prior audit)	0.012	0.012	0.012	0.010
	[0.022]	[0.022]	[0.022]	[0.022]
Paid by unknown entity (focal audit)	-0.022	-0.019	-0.020	-0.025
	[0.067]	[0.065]	[0.066]	[0.071]
Prior audit was re-audit, but focal audit was not	-0.120**	-0.123**	-0.124**	-0.123**
	[0.035]	[0.033]	[0.033]	[0.033]
Focal audit was re-audit, but prior audit was not	0.279**	0.277**	0.276**	0.276**
	[0.016]	[0.016]	[0.017]	[0.017]
Prior and focal audit were re-audits	0.198**	0.195**	0.194**	0.195**
	[0.011]	[0.013]	[0.013]	[0.013]
Audit duration† (prior audit)	0.125**	0.131**	0.128**	0.130**
	[0.026]	[0.025]	[0.024]	[0.025]
Previous auditor (prior audit)	0.012	0.014	0.015	0.014
	[0.026]	[0.026]	[0.026]	[0.026]
Previous auditor (focal audit)	0.032+	0.032+	0.033+	0.032+
	[0.018]	[0.018]	[0.018]	[0.018]

(continued on next page)

Table 4 (continued)

	(1)	(2)	(3)	(4)
All-female audit team (prior audit)	0.009	0.009	0.008	0.008
	[0.013]	[0.013]	[0.014]	[0.014]
Mixed-gender audit team (prior audit)	-0.021	-0.021	-0.020	-0.021
	[0.025]	[0.025]	[0.024]	[0.024]
All-female audit team (focal audit)	-0.064**	-0.063**	-0.064**	-0.063**
	[0.018]	[0.018]	[0.018]	[0.018]
Mixed-gender audit team (focal audit)	-0.122**	-0.122**	-0.122**	-0.122**
	[0.027]	[0.027]	[0.027]	[0.027]
Maximum auditor tenure† (prior audit)	-0.184	-0.177	-0.175	-0.174
	[0.197]	[0.193]	[0.193]	[0.192]
Maximum auditor tenure† (prior audit), squared	0.048	0.045	0.045	0.044
	[0.051]	[0.050]	[0.050]	[0.050]
Maximum auditor tenure† (focal audit)	-0.562**	-0.571**	-0.573**	-0.564**
	[0.170]	[0.170]	[0.171]	[0.174]
Maximum auditor tenure† (focal audit), squared	0.151**	0.153**	0.154**	0.152**
	[0.042]	[0.042]	[0.043]	[0.043]
FDI inflows† (prior audit)	-0.065	-0.065	-0.064	-0.065
	[0.040]	[0.040]	[0.040]	[0.040]
Labor laws† (2002)	0.160	0.161	0.159	0.159
	[0.115]	[0.116]	[0.116]	[0.116]
GDP per capita† (prior audit)	0.052**	0.054**	0.055**	0.055**
	[0.016]	[0.016]	[0.016]	[0.016]
GDP per capita in buyer country† (prior audit)	-0.005	0.005	0.006	0.005
	[0.029]	[0.031]	[0.031]	[0.031]
Buyer employment† (prior year)	-0.040**	-0.041**	-0.042**	-0.042**
	[0.005]	[0.005]	[0.005]	[0.005]
Year fixed effects	Included	Included	Included	Included
Sequence fixed effects	Included	Included	Included	Included
Industry fixed effects	Included	Included	Included	Included
Buyer-country fixed effects	Included	Included	Included	Included
R-squared	0.41	0.41	0.41	0.42
F value of Wald test: are the two interaction coefficients equivalent?			22.08**	13.39**

Notes. Ordinary least squares (OLS) regression coefficients with standard errors clustered by supplier country in brackets. N = 8,677 audits of 4,940 factories. **p<0.01, *p<0.05, +p<0.10. † indicates logged. § indicates logged, then standardized. Baseline (omitted) categories are *paid by the buyer* for focal and prior audit and that neither the focal nor prior are re-audits. All models include dummy variables to indicate instances in which the following variables were missing data and thus recoded to 0: *buyer exposure sensitivity (prior audit)* (903 audits), *audit duration (prior audit)* (4,339 audits), *previous auditor (prior audit)* (4,154 audits), *previous auditor (focal audit)* (9 audits), and *buyer employment* (3,322 audits).

Table 5. Extension Results

Dependent variable:	(1) Child labor score improvement	(2) Working hours score improvement	(3) Minimum wage score improvement	(4) OSH score improvement
Exposure risk (prior audit)	0.044**	0.034**	0.059**	0.043*
	[0.006]	[0.012]	[0.010]	[0.017]
Buyer exposure sensitivity (prior audit)	0.008	0.038**	0.030*	0.020
	[0.013]	[0.012]	[0.013]	[0.016]
Announced (prior audit)	-0.014**	0.009	0.009	0.039**
	[0.004]	[0.010]	[0.005]	[0.011]
Maximum auditor training§ (prior audit)	0.017*	0.010	0.034**	0.024*
	[0.008]	[0.009]	[0.006]	[0.012]
Announced (focal audit)	0.036**	0.014	0.053**	0.048**
	[0.012]	[0.017]	[0.012]	[0.016]
Maximum auditor training† (focal audit)	-0.004	-0.044**	-0.015+	-0.044+
	[0.005]	[0.010]	[0.008]	[0.023]
Number of child labor violations (prior audit)	0.383**			
	[0.005]			
Number of hours of work violations (prior audit)		0.315**		
		[0.007]		
Number of minimum wage violations (prior audit)			0.249**	
			[0.010]	
Number of OSH violations (prior audit)				0.121**
				[0.008]
Paid by supplier or third party (prior audit)	-0.002	0.023**	-0.044+	-0.039*
	[0.006]	[0.007]	[0.023]	[0.015]
Paid by supplier or third party (focal audit)	0.027**	0.050**	0.081**	0.024
	[0.007]	[0.015]	[0.021]	[0.019]
Paid by unknown entity (prior audit)	0.001	0.013	0.006	0.053*
	[0.006]	[0.028]	[0.016]	[0.024]
Paid by unknown entity (focal audit)	0.007	-0.020	-0.016	-0.018
	[0.010]	[0.038]	[0.039]	[0.053]
Prior audit was re-audit, but focal audit was not	-0.009	-0.049**	-0.054**	-0.191**
	[0.011]	[0.013]	[0.010]	[0.054]
Focal audit was re-audit, but prior audit was not	0.030**	0.030**	0.073**	0.307**
	[0.007]	[0.009]	[0.014]	[0.025]
Prior and focal audit were re-audits	0.037**	0.006	0.041**	0.245**
	[0.005]	[0.009]	[0.013]	[0.010]
Audit duration† (prior audit)	0.051**	0.090**	0.087**	0.168**
	[0.016]	[0.018]	[0.014]	[0.039]
Previous auditor (prior audit)	0.007	-0.001	-0.012	0.020
	[0.005]	[0.017]	[0.017]	[0.026]
Previous auditor (focal audit)	-0.011	-0.014	0.017	0.037*
	[0.013]	[0.014]	[0.021]	[0.017]

(continued on next page)

Table 5 (continued)

All-female audit team (prior audit)	0.021**	0.002	-0.009	0.007
	[0.007]	[0.010]	[0.014]	[0.013]
Mixed-gender audit team (prior audit)	0.018*	-0.012	-0.026	0.005
	[0.008]	[0.011]	[0.020]	[0.035]
All-female audit team (focal audit)	-0.004	-0.035*	-0.065**	-0.045*
	[0.009]	[0.015]	[0.011]	[0.020]
Mixed-gender audit team (focal audit)	0.011	-0.051*	-0.067**	-0.128**
	[0.014]	[0.020]	[0.018]	[0.030]
Maximum auditor tenure† (prior audit)	0.051	0.125	0.062	0.287
	[0.065]	[0.095]	[0.095]	[0.336]
Maximum auditor tenure† (prior audit), squared	-0.007	-0.039	-0.013	-0.075
	[0.017]	[0.028]	[0.028]	[0.092]
Maximum auditor tenure† (focal audit)	-0.172**	-0.262*	-0.139	-0.396*
	[0.050]	[0.113]	[0.102]	[0.175]
Maximum auditor tenure† (focal audit), squared	0.042**	0.060*	0.035	0.121*
	[0.014]	[0.030]	[0.031]	[0.046]
FDI inflows† (prior audit)	-0.026*	-0.068*	-0.031	-0.024
	[0.013]	[0.028]	[0.022]	[0.042]
Labor laws† (2002)	0.011	0.184	0.021	-0.004
	[0.046]	[0.123]	[0.072]	[0.101]
GDP per capita† (prior audit)	0.001	0.045**	0.035**	0.031+
	[0.006]	[0.015]	[0.010]	[0.018]
GDP per capita in buyer country† (prior audit)	0.015	-0.042	0.032+	0.025
	[0.010]	[0.025]	[0.019]	[0.020]
Buyer employment† (prior year)	-0.008**	-0.021**	-0.002	-0.019**
	[0.002]	[0.004]	[0.004]	[0.007]
Year fixed effects	Included	Included	Included	Included
Sequence fixed effects	Included	Included	Included	Included
Industry fixed effects	Included	Included	Included	Included
Buyer-country fixed effects	Included	Included	Included	Included
R-squared	0.12	0.14	0.17	0.33
Mean y in sample	0.07	0.05	0.08	0.25

Notes. OLS regression coefficients with standard errors clustered by supplier country in brackets. N =8,335 audits of 4,870 factories, slightly smaller than in Table 4 due to a few missing values of violation category counts. **p<0.01, *p<0.05, +p<0.10. † indicates logged. § indicates logged, then standardized. See Table 4 for additional notes.

Appendix

Table A1. Correlations

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)
(1) Improvement	1.00																
(2) Exposure risk (prior audit)	-0.03	1.00															
(3) Press freedom§ (prior audit)	-0.03	0.96	1.00														
(4) NGO density§ (prior audit)	-0.02	0.95	0.84	1.00													
(5) Buyer exposure sensitivity (prior audit)	0.02	0.12	0.11	0.11	1.00												
(6) Announced (prior audit)	0.07	-0.13	-0.15	-0.10	-0.04	1.00											
(7) Maximum auditor training§ (prior audit)	0.06	-0.17	-0.18	-0.15	-0.08	0.02	1.00										
(8) Violations† (prior audit)	0.57	-0.29	-0.29	-0.26	0.01	0.09	0.04	1.00									
(9) Announced (focal audit)	0.05	-0.17	-0.19	-0.13	-0.06	0.49	0.02	0.06	1.00								
(10) Maximum auditor training† (focal audit)	-0.02	-0.18	-0.17	-0.18	-0.08	0.06	0.69	0.04	0.06	1.00							
(11) Paid by supplier or third party (prior audit)	0.00	-0.28	-0.29	-0.23	-0.33	0.16	-0.03	0.05	0.19	-0.04	1.00						
(12) Paid by the buyer (prior audit)	-0.02	0.26	0.27	0.22	0.32	-0.18	0.03	-0.07	-0.20	0.02	-0.88	1.00					
(13) Paid by unknown entity (prior audit)	0.04	0.05	0.05	0.04	0.04	0.04	0.02	0.05	0.01	0.04	-0.26	-0.22	1.00				
(14) Paid by supplier or third party (focal audit)	0.04	-0.26	-0.28	-0.22	-0.40	0.13	-0.03	0.06	0.23	-0.04	0.75	-0.70	-0.10	1.00			
(15) Paid by the buyer (focal audit)	-0.05	0.26	0.27	0.21	0.39	-0.14	0.02	-0.09	-0.24	0.03	-0.70	0.77	-0.12	-0.90	1.00		
(16) Paid by unknown entity (focal audit)	0.01	0.02	0.03	0.01	0.03	0.01	0.02	0.06	0.01	0.03	-0.12	-0.12	0.50	-0.25	-0.20	1.00	
(17) Prior audit was re-audit, but focal audit was not	-0.21	-0.03	-0.03	-0.03	-0.05	-0.02	-0.08	-0.18	0.06	0.03	0.03	-0.02	-0.03	0.01	-0.01	-0.01	1.00
(18) Focal audit was re-audit, but prior audit was not	0.27	-0.08	-0.08	-0.07	0.04	0.07	0.05	0.31	-0.02	-0.06	0.05	-0.07	0.04	0.04	-0.05	0.03	-0.26
(19) Prior and focal audit were re-audits	0.02	-0.15	-0.16	-0.13	0.00	-0.02	0.04	-0.02	0.00	-0.04	0.12	-0.09	-0.06	0.11	-0.08	-0.06	-0.17
(20) Audit duration† (prior audit)	0.22	-0.39	-0.41	-0.32	0.07	0.19	0.24	0.37	0.22	0.26	0.04	-0.05	0.02	0.05	-0.06	0.02	-0.03
(21) Previous auditor (prior audit)	0.01	0.29	0.28	0.28	0.05	-0.08	-0.03	-0.07	-0.09	-0.06	-0.11	0.10	0.02	-0.11	0.11	0.00	-0.02
(22) Previous auditor (focal audit)	0.03	0.34	0.33	0.32	0.04	-0.04	-0.04	-0.08	-0.07	-0.09	-0.12	0.10	0.04	-0.10	0.10	0.02	-0.04
(23) Audit sequence	-0.08	-0.08	-0.08	-0.08	-0.05	-0.02	0.29	-0.17	-0.01	0.23	0.04	-0.01	-0.07	0.03	0.00	-0.06	0.16
(24) All-female audit team (prior audit)	0.00	-0.13	-0.12	-0.13	0.00	0.03	-0.02	0.02	0.03	0.01	0.03	-0.03	-0.01	0.03	-0.03	0.00	0.02
(25) Mixed-gender audit team (prior audit)	0.02	-0.01	-0.01	-0.02	-0.02	0.00	0.09	0.03	0.00	0.05	0.01	-0.01	0.01	0.00	-0.01	0.01	-0.01
(26) All-female audit team (focal audit)	-0.01	-0.16	-0.16	-0.15	-0.03	0.00	0.00	0.02	0.02	0.00	0.05	-0.05	-0.01	0.05	-0.05	-0.01	0.01
(27) Mixed-gender audit team (focal audit)	-0.04	0.01	0.02	0.00	-0.02	0.01	0.01	0.00	0.02	0.05	0.01	-0.01	0.00	0.01	-0.01	0.00	0.03
(28) Maximum auditor tenure† (prior audit)	-0.01	0.20	0.20	0.17	0.03	-0.07	-0.34	-0.02	-0.07	-0.32	0.00	-0.01	0.03	0.00	0.00	0.02	0.02
(29) Maximum auditor tenure† (focal audit)	-0.01	0.19	0.19	0.17	0.01	-0.06	-0.32	-0.03	-0.08	-0.27	0.02	-0.02	0.00	0.02	-0.03	0.03	-0.02
(30) Buyer employment † (prior year)	0.03	-0.25	-0.27	-0.20	-0.19	0.25	0.04	0.09	0.33	0.06	0.63	-0.65	0.03	0.71	-0.72	0.02	0.03
(31) FDI inflows† (prior audit)	0.03	-0.44	-0.55	-0.28	-0.07	0.11	0.27	0.16	0.15	0.27	0.18	-0.17	-0.01	0.18	-0.18	-0.01	0.01
(32) Labor laws† (2002)	0.02	-0.43	-0.41	-0.41	-0.05	0.03	0.10	0.09	0.06	0.13	0.09	-0.09	0.00	0.10	-0.11	0.02	0.00
(33) GDP per capita† (prior audit)	0.00	0.58	0.59	0.53	0.08	-0.16	0.05	-0.20	-0.17	0.03	-0.21	0.19	0.04	-0.21	0.19	0.04	-0.03
(34) GDP per capita in buyer country† (prior audit)	0.01	0.03	0.04	0.02	0.03	0.12	0.11	-0.06	0.08	0.09	0.00	0.13	-0.25	-0.05	0.10	-0.11	-0.01

	(18)	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)	(28)	(29)	(30)	(31)	(32)	(33)	(34)
(18) Focal audit was re-audit, but prior audit was not	1.00																
(19) Prior and focal audit were re-audits	-0.38	1.00															
(20) Audit duration† (prior audit)	0.12	0.00	1.00														
(21) Previous auditor (prior audit)	-0.03	-0.02	-0.07	1.00													
(22) Previous auditor (focal audit)	0.03	-0.04	-0.07	0.23	1.00												
(23) Audit sequence	-0.24	0.26	-0.04	0.01	-0.03	1.00											
(24) All-female audit team (prior audit)	-0.01	0.02	-0.06	-0.11	-0.12	0.01	1.00										
(25) Mixed-gender audit team (prior audit)	0.00	0.00	0.23	0.04	0.07	0.06	-0.47	1.00									
(26) All-female audit team (focal audit)	0.01	0.03	-0.03	-0.10	-0.10	0.02	0.14	-0.04	1.00								
(27) Mixed-gender audit team (focal audit)	-0.02	-0.01	0.07	0.00	0.03	0.03	-0.02	0.16	-0.47	1.00							
(28) Maximum auditor tenure† (prior audit)	-0.04	-0.02	-0.18	0.06	0.04	-0.09	0.02	0.04	0.01	0.00	1.00						
(29) Maximum auditor tenure† (focal audit)	0.01	0.02	-0.21	0.08	0.09	-0.07	0.02	-0.03	0.04	0.02	0.36	1.00					
(30) Buyer employment † (prior year)	0.02	0.07	0.14	-0.14	-0.12	0.06	0.03	0.00	0.05	0.00	-0.02	-0.01	1.00				
(31) FDI inflows† (prior audit)	0.05	0.08	0.28	-0.18	-0.22	0.11	0.05	0.02	0.09	0.00	-0.19	-0.18	0.21	1.00			
(32) Labor laws† (2002)	-0.01	0.05	0.08	-0.23	-0.25	0.08	0.20	0.00	0.21	-0.01	-0.03	-0.04	0.13	0.25	1.00		
(33) GDP per capita† (prior audit)	-0.09	-0.10	-0.34	0.12	0.13	0.10	0.06	-0.04	0.05	-0.02	0.22	0.23	-0.14	-0.21	0.10	1.00	
(34) GDP per capita in buyer country† (prior audit)	0.00	0.01	0.04	0.00	-0.01	0.09	0.01	0.01	0.01	0.00	-0.03	-0.01	0.03	-0.01	0.02	0.09	1.00

Notes. † indicates logged. § indicates logged, then standardized. Piecewise correlations; see Table 3 footer for number of observations per variable.

Table A2. Negative Binomial Regression Results

Dependent variable: Number of violations

	(1)	(2)	(3)	(4)
Exposure risk (prior audit)	-0.130**	-0.130**	-0.126**	-0.129**
	[0.025]	[0.025]	[0.026]	[0.025]
Buyer exposure sensitivity (prior audit)	-0.066**	-0.070**	-0.070**	-0.072**
	[0.022]	[0.021]	[0.021]	[0.021]
Announced (prior audit)	-0.034*	-0.033*	-0.032*	-0.032*
	[0.014]	[0.015]	[0.015]	[0.015]
Maximum auditor training§ (prior audit)	-0.029*	0.007	0.012	-0.000
	[0.013]	[0.021]	[0.018]	[0.020]
Maximum auditor training§ (prior audit) × Announced (prior audit)		-0.046**	-0.043**	-0.046**
		[0.014]	[0.013]	[0.015]
Maximum auditor training§ (prior audit) × Exposure risk (prior audit)			0.011	
			[0.010]	
Maximum auditor training§ (prior audit) × Buyer exposure sensitivity (prior audit)				0.020
				[0.015]
Announced (focal audit)	-0.106**	-0.103**	-0.102**	-0.101**
	[0.017]	[0.018]	[0.017]	[0.018]
Maximum auditor training† (focal audit)	0.047	0.047	0.048	0.047
	[0.030]	[0.030]	[0.030]	[0.030]
Violations† (prior audit)	0.333**	0.330**	0.330**	0.330**
	[0.020]	[0.021]	[0.021]	[0.021]
Paid by supplier or third party (prior audit)	0.025	0.028	0.028	0.029
	[0.026]	[0.025]	[0.025]	[0.025]
Paid by supplier or third party (focal audit)	-0.115**	-0.120**	-0.120**	-0.123**
	[0.022]	[0.022]	[0.022]	[0.021]
Paid by unknown entity (prior audit)	-0.007	-0.007	-0.008	-0.007
	[0.031]	[0.031]	[0.030]	[0.031]
Paid by unknown entity (focal audit)	0.021	0.019	0.020	0.027
	[0.070]	[0.069]	[0.070]	[0.074]
Prior audit was re-audit, but focal audit was not	0.134**	0.136**	0.137**	0.136**
	[0.039]	[0.038]	[0.038]	[0.038]
Focal audit was re-audit, but prior audit was not	-0.255**	-0.254**	-0.253**	-0.253**
	[0.014]	[0.014]	[0.015]	[0.015]
Prior and focal audit were re-audits	-0.190**	-0.187**	-0.185**	-0.187**
	[0.014]	[0.015]	[0.015]	[0.015]
Audit duration† (prior audit)	-0.172**	-0.177**	-0.171**	-0.177**
	[0.035]	[0.035]	[0.032]	[0.034]
Previous auditor (prior audit)	-0.018	-0.018	-0.021	-0.019
	[0.033]	[0.033]	[0.033]	[0.032]
Previous auditor (focal audit)	-0.032	-0.032	-0.032	-0.032
	[0.021]	[0.021]	[0.021]	[0.021]

(continued on next page)

Table A2 (continued)

All-female audit team (prior audit)	-0.008	-0.008	-0.007	-0.007
	[0.018]	[0.018]	[0.019]	[0.018]
Mixed-gender audit team (prior audit)	0.018	0.017	0.016	0.018
	[0.036]	[0.036]	[0.035]	[0.035]
All-female audit team (focal audit)	0.077**	0.077**	0.077**	0.077**
	[0.026]	[0.026]	[0.026]	[0.027]
Mixed-gender audit team (focal audit)	0.156**	0.156**	0.155**	0.156**
	[0.042]	[0.042]	[0.042]	[0.043]
Maximum auditor tenure† (prior audit)	-0.168	-0.175	-0.174	-0.184
	[0.218]	[0.215]	[0.213]	[0.212]
Maximum auditor tenure† (prior audit), squared	0.038	0.041	0.041	0.044
	[0.059]	[0.059]	[0.058]	[0.058]
Maximum auditor tenure† (focal audit)	0.864**	0.873**	0.877**	0.870**
	[0.180]	[0.180]	[0.180]	[0.181]
Maximum auditor tenure† (focal audit), squared	-0.230**	-0.233**	-0.234**	-0.232**
	[0.046]	[0.046]	[0.046]	[0.046]
FDI inflows† (prior audit)	0.085	0.085	0.083	0.084
	[0.067]	[0.067]	[0.066]	[0.067]
Labor laws† (2002)	-0.251	-0.251	-0.246	-0.249
	[0.200]	[0.201]	[0.200]	[0.201]
GDP per capita† (prior audit)	-0.072**	-0.075**	-0.076**	-0.076**
	[0.027]	[0.027]	[0.027]	[0.027]
GDP per capita in buyer country† (prior audit)	-0.029	-0.040+	-0.040+	-0.039+
	[0.020]	[0.021]	[0.021]	[0.021]
Buyer employment† (prior year)	0.027**	0.028**	0.029**	0.029**
	[0.006]	[0.007]	[0.006]	[0.006]
Year fixed effects	Included	Included	Included	Included
Sequence fixed effects	Included	Included	Included	Included
Industry fixed effects	Included	Included	Included	Included
Buyer-country fixed effects	Included	Included	Included	Included
Chi-squared value of Wald test: are the two interaction coefficients equivalent?			13.64**	7.394**

Notes. Negative binomial regression coefficients with standard errors clustered by supplier country in brackets. Average number of violations is 5.64. N = 8,677 audits of 4,940 factories. **p<0.01, *p<0.05, +p<0.10. † indicates logged. § indicates logged, then standardized. Baseline (omitted) categories are *paid by the buyer* for focal and prior audit and that the neither focal nor prior are re-audits. All models include dummy variables to indicate instances in which the following variables were missing data and thus recoded to 0: *buyer exposure sensitivity (prior audit)*, *audit duration (prior audit)*, *previous auditor (prior audit)*, *previous auditor (focal audit)*, and *buyer employment*. Because the dependent variable in these models is the number of violations, negatively signed coefficients indicate greater improvement (that is, fewer violations), in contrast to our primary models, in which positively signed coefficients refer to greater improvement.