


IMPROVING WORKING CONDITIONS IN GLOBAL SUPPLY CHAINS: THE ROLE OF INSTITUTIONAL ENVIRONMENTS AND MONITORING PROGRAM DESIGN

JODI L. SHORT, MICHAEL W. TOFFEL, AND ANDREA R. HUGILL*

Activism seeking to improve labor conditions in global supply chains has led many transnational corporations to adopt codes of conduct and to monitor suppliers for compliance. Drawing on thousands of audits conducted by a major social auditor, the authors identify structural contingencies in the institutional environment and program design under which codes and monitoring are more likely to be associated with improvements in conditions. At the institutional level, suppliers improve more when they face greater risk that nongovernmental organizations and the press will expose harmful working conditions. They also improve more when their buyers have experienced negative publicity for supply chain labor abuses. At the program design level, suppliers improve more on average when audits are pre-announced, when auditors are highly trained, and especially when both elements are present. Extended analysis of variations across violation types reveals nuances to these findings. For instance, pre-announced audits were followed by greater improvement in occupational safety and health practices but not child labor practices. These findings can inform strategies for improving supply chain working conditions.

*JODI L. SHORT ( <https://orcid.org/0000-0002-0061-1966>) is the Honorable Roger J. Traynor Professor of Law at the UC Hastings College of the Law. MICHAEL W. TOFFEL is the Senator John Heinz Professor of Environmental Management at the Harvard Business School. ANDREA R. HUGILL is an Adjunct Professor at the McDonough School of Business, Georgetown University.

We gratefully acknowledge the research assistance provided by Melissa Ouellet as well as by Chris Allen, John Galvin, Erika McCaffrey, and Christine Rivera. We appreciate the helpful comments provided by Xiang Ao, Andrew Marder, Bill Simpson, Benjamin van Rooij, and the participants of the University of Michigan Ross School of Business strategy seminar, the Ohio State University Fisher College of Business management sciences seminar, the MIT Sloan School of Management operations management seminar, the UC Irvine Long US–China Institute workshop, the Tulane Law School Murphy Institute workshop series, the Duke Strategy/EDGE Center Seminar, and the University of Texas–Austin McCombs School of Business operations management seminar. We appreciate the financial support provided by Harvard Business School’s Division of Research and Faculty Development. An Online Appendix is available at <http://journals.sagepub.com/doi/suppl/10.1177/0019793920916181>. For information regarding the data and/or computer programs used for this study, please address correspondence to shortj@uchastings.edu.

KEYWORDS: codes of conduct, supply chain labor standards, compliance, decoupling, private politics

Widely reported labor abuses and pressure from consumers and labor activists have prompted high-profile brands to adopt formal organizational structures, such as codes of labor conduct and social monitoring, to address risks in their global supply chains. Nike developed an extensive program of codes and monitoring following reports of child labor in its supplier factories; Apple strengthened its supplier monitoring program in response to worker suicides and reports of rampant labor abuses at its supplier Foxconn's factory in China; and prominent European apparel brands enhanced their existing programs following the collapse of the Rana Plaza building that housed their suppliers, killing more than 1,000 Bangladeshi factory workers.

Beyond such high-profile cases, codes and monitoring have become ubiquitous in global value chains. To avoid negative publicity generated by accidents and activism, thousands of transnational corporations (TNCs), including all US Fortune 500 companies, have adopted codes of conduct that require their suppliers to meet specified workplace standards (McBarnet 2007), and many conduct social audits to monitor and assess suppliers' adherence to those codes (Short, Toffel, and Hugill 2016). Codes and monitoring are also used by multi-stakeholder initiatives—such as the Roundtable on Sustainable Palm Oil, the Electronics Industry Citizenship Coalition, and the Ethical Trading Initiative—that provide collective fora for private regulation of supply chain practices. Hundreds of thousands of audits are conducted on behalf of TNCs and multi-stakeholder initiatives each year (Gould 2005). Some activist groups have claimed that social auditing is an \$80 billion/year industry (AFL-CIO 2013).

While TNCs generally adopt 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 strategy to improve conditions for 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 monitoring. United Nations initiatives in the area of business and human rights, such as the Global Compact and the Guiding Principles on Business and Human Rights, rely on private standards and monitoring to improve companies' global labor and other human rights practices (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” (UL Responsible Sourcing 2015 [emphasis added]; see also Elevate 2016; Intertek 2016).

However, it is not clear when, or even whether, codes and monitoring improve supplier labor practices. Many observers have argued that such organizational structures are, at best, window dressing (Esbenshade 2004;

Frynas 2005; Seidman 2007; Barkemeyer 2009) or, worse, “organized hypocrisy” (Lim and Tsutsui 2012: 69)—that is, 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 symbolic structures that are likely to be decoupled from actual supplier practices (Meyer and Rowan 1977).

Nevertheless, we argue that decoupling is not inevitable. We extend the literature on 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 ask: How do suppliers’ compliance improvement rates differ depending on variations in institutional compliance pressures and in the design of monitoring programs? We also examine the interaction among program design elements. Identifying the conditions under which codes and monitoring are associated with actual improvements in labor conditions is critically important because they are the dominant mode of labor standards regulation in global value chains. We make no claims about their efficacy relative to other vehicles for improving labor standards, including more robust state-based regulation or worker organizing and empowerment. Rather, our study explores how the efficacy of the existing, highly institutionalized system of codes and monitoring could be enhanced.

We test our hypotheses using a novel data set of thousands of audits for code-of-conduct compliance conducted at nearly 5,000 factories spanning 13 industries in 66 countries by one of the world’s largest supply chain auditing firms. We use regression analysis to estimate a model that predicts a factory’s improvement in code compliance based on our hypothesized institutional pressures and program design attributes, controlling for audit and institution characteristics that might also affect improvement. By identifying structural contingencies that favor coupling, our findings challenge the assumption that codes and monitoring will inevitably be decoupled from practices and suggest key considerations that should inform the design and implementation of monitoring strategies to improve conditions in global supply chains.

Decoupling of Labor Codes and Labor Practices

Meyer and Rowan (1977) defined formal organizational structures as sets of practices and procedures that embody rationalized concepts of organizational legitimacy and theorized the conditions under which such structures would be decoupled from actual organizational practice. The decoupling literature strongly suggests that formal organizational structures like labor codes of conduct and social monitoring are likely to be adopted symbolically and decoupled from the suppliers’ actual labor practices.

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 be implemented symbolically and decoupled from practices (Meyer and Rowan 1977; Boiral 2007; Bromley and Powell 2012). 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: 341) than codes and monitoring. Companies initially adopted codes and monitoring to deflect negative publicity and preserve brand reputation (Bartley and Child 2014). However, some of the substantive changes in production practices that such structures demand—minimum wage requirements, overtime restrictions, and freedom of association rights—are fundamentally at odds with the economic logic of global value chains that seek to minimize production costs (Locke 2013).

Second, symbolic structures are more likely to be decoupled in contexts in which efficiency demands are strong and not tempered by countervailing institutional pressures (Meyer and Rowan 1977; Meyer, Boli, Thomas, and Ramirez 1997). Suppliers to global value chains face intense efficiency demands to produce high volume at low cost (Gereffi, Humphrey, and Sturgeon 2005), and many are in countries with weak regulatory institutions and lax enforcement of labor standards.

Finally, resource constraints impede substantive implementation of formal organizational structures (Meyer et al. 1997; Bromley and Powell 2012; Lim and Tsutsui 2012). Suppliers operate on razor-thin margins and many lack the resources to effectively implement code requirements. Thus, the axioms that emerge from the decoupling literature suggest that formal 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” (Bromley and Powell 2012: 489) in ways that improve workers’ conditions.

Against this grain, a growing research stream focuses on the conditions under which formal organizational structures adopted symbolically are nevertheless implemented substantively or are coupled with organizational practices (Bromley and Powell 2012; Bartley and Egels-Zandén 2015). Consistent with the decoupling literature, most studies that 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 anti-discrimination law (Kalev, Dobbin, and Kelly 2006) and environmental law (Potoski and Prakash 2005).

Studies specifically investigating suppliers’ compliance with labor codes of conduct similarly find that codes and monitoring are associated with better working conditions when combined with robust 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; Distelhorst, Hainmueller, and Locke 2017). In addition to the coercive power of the state, studies have found that institutional pressures from unions (Bartley and Egels-Zandén 2015; Oka 2015), a free press (Toffel, Short, and Ouellet 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 with their actual labor practices.

Some qualitative studies have expanded the decoupling literature's traditional focus on coercive institutions to investigate how the activities of individual actors inside organizations can create contingencies that promote coupling (Espeland 1998; Hallett 2010; Overdevest 2010; Tilcsik 2010; Bartley and Egels-Zandén 2015). For example, in a study of Indonesian apparel and footwear factories, Bartley and Egels-Zandén (2015) found that the coupling of labor codes of conduct and supplier labor practices was contingent on union members' ability to forge and leverage relationships with brands, international NGOs, and global unions to pressure suppliers to meet their code commitments.

Bartley and Egels-Zandén (2015) highlighted the importance of identifying contingencies that can promote coupling in contexts where theory conventionally predicts a high likelihood of decoupling. To date, such research has focused largely on the "thorny, on-the-ground processes" (Bartley and Egels-Zandén 2016: 233) that couple organizational structures and practices 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). We extend this research with a large-scale study that examines structural contingencies of coupling across a range of institutional contexts. Our hypotheses investigate the coupling potential of structural contingencies 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 1) institutional pressure from civil society groups on suppliers and brands and 2) key design features of the monitoring programs that brands adopt. We focus on these factors because they are structural contingencies over which brands and/or activists might exercise influence. Brands, for instance, must choose how to design their auditing programs, and activists must choose where to target their limited resources and what types of pressure to apply. Our hypothesized variables yield insights that can inform these decisions.

Our work builds on existing large-scale studies of coupling and labor codes of conduct in three important ways. First, we examine contingencies at the level of program design that other studies have overlooked (e.g., Toffel, Short, and Ouellet 2015; Bird, Short, and Toffel 2019) or have rejected as explaining compliance outcomes. Indeed, recent studies that explain coupling outcomes based on local institutions (Distelhorst, Locke, Pal, and Samel 2015) or social structures governing labor relations (Distelhorst et al. 2017) explicitly note that monitoring system design is not

driving these outcomes. This gap in the literature mirrors the broader literature on decoupling where, with few exceptions (e.g., Kalev, Dobbin, and Kelly 2006), little attention has been paid to the influence of program design on coupling outcomes. Second, we examine and find contingencies operating at multiple levels: from the details of program design to broad institutional pressures. Existing studies tend to focus on one level—for instance, either institutional (Toffel, Short, and Ouellet 2015) or organizational (Distelhorst et al. 2017; Bird, Short, and Toffel 2019)—or they fail to identify hypothesized contingencies at multiple levels (Distelhorst et al. 2015).

Finally, we measure coupling by the *improvement* in suppliers' compliance with the labor standards contained in codes of conduct. Most research in this field has measured levels of supplier compliance at a moment in time rather than improvement over time (e.g., Egels-Zandén 2007; Locke, Qin, and Brause 2007; Oka 2010a, 2010b; Ang, Brown, Dehejia, and Robertson 2012; Bartley and Egels-Zandén 2015; Toffel, Short, and Ouellet 2015). The extent to which suppliers change their practices over time to conform to codes of conduct is a critical marker of whether codes and practices are coupled within supplier organizations. In our empirical context, however, little is known about why some suppliers improve their compliance more than others. 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, Lund-Thomsen, Xue, and Khara 2011; Ang et al. 2012; Locke, Rissing, and Pal 2013; Toffel, Short, and Ouellet 2015). However, these studies have not hypothesized conditions associated with improvement.

We build on a handful of studies that have examined empirically the factors associated with code compliance improvement. Some measure improvement at the national or regional level, leaving open the question of what accounts for variation in individual suppliers' improvement (Weil and Mallo 2007; Ang et al. 2012). Others identify characteristics shared among small samples of firms that all improved their code compliance, but these studies fail to identify factors that explain why some factories improved more than others (e.g., Egels-Zandén 2007; Locke, Rissing, and Pal 2013). A few studies have revealed variation in improvement across factories based on the internal management structures in supplier organizations (Oka 2015; Distelhorst et al. 2017; Bird et al. 2019), but they do not look at influences outside the organization. Our study contributes to this literature by identifying both institutional and programmatic contingencies that explain variation in supplier improvement.

Institutional Compliance Pressures

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” (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). We explore two distinct but related sources of institutional pressure that increase reputational risk for global buyers: 1) the ability of civil society actors to discover and expose supplier abuses and 2) buyers’ past exposure to negative publicity about their suppliers’ practices.

Institutional Pressure in Suppliers’ Domestic Environments

Although government inspection regimes are often weak in countries where suppliers are located, research suggests that civil society actors, such as NGOs and the press, can provide monitoring functions and expose wrongdoing (Seidman 2007; 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 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).

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 et al. (2015), for instance, demonstrated that suppliers in countries with more press freedom exhibit greater compliance with codes of conduct. However, it is not clear that the mechanisms fostering high compliance *levels* will also foster *improvements* in compliance. Indeed, in high-compliance environments, there may be less room to improve because of a lack of low-hanging fruit (Chatterji and Toffel 2010), or because there may simply be less pressure to improve if conditions are not so bad. 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 is related to compliance levels. For instance, domestic environments with very low or very high labor-standards compliance might attract more NGOs,

but a distinct and important question is whether their presence will be associated with *improvements* in compliance.

We argue that institutional pressure generated by civil society actors such as the press and NGOs, and particularly those two in combination, will be associated with improvements in supplier compliance. Suppliers whose failure to meet the global norms prescribed by codes of conduct is documented in audits become attractive targets for transnational advocacy networks seeking to raise international labor and human rights standards.¹ These networks seek 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 labor abuses at suppliers, increasing the prospect that the suppliers—and others like them—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 their labor practices more when located in institutional environments in which there is greater potential for civil society monitoring mechanisms to expose noncompliance with codes of conduct.

Targeted Institutional Pressure on Buyers

The impact of the institutional pressures described above depend not only on the probability that suppliers’ wrongdoing will be exposed but also on buyers’ reaction to such revelations. The reputational stakes of exposure are higher for some buyers than for others. 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 highly reputation-conscious buyers are more likely to work with suppliers that better comply with labor standards (Oka 2010a).

Less is known, however, about how buyers that have experienced negative reputation events might respond to the threat of additional reputational

¹In a very different context, studies have suggested that activists and the press are less likely to target the worst-behaved companies than companies that already have strong reputations for social responsibility performance (Luo, Meier, and Oberholzer-Gee 2012; Bartley and Child 2014) or companies that have already adopted extensive organizational structures to implement their corporate social responsibility (CSR) initiatives (McDonnell, King, and Soule 2015). 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 because they face greater financial consequences of reputational damage. Such is not the case with suppliers in developing countries. They are unbranded, largely invisible to consumers, and thus more insulated from reputational threat. We believe that activists will select their targets very differently in these contexts and will attempt to identify the worst practices by local suppliers in order to gain the most leverage over their global brand targets.

damage. It has been theorized that firms facing the prospect of reputational threats from activists, as many multinational buyers do, will attempt to forestall trouble by investing in self-regulatory activities (Abito et al. 2016). Buyers that have already suffered reputational shocks through criticism by or confrontation with activists are especially likely to invest in self-regulatory measures to avoid additional reputational harm (Abito et al. 2016). Because a damaged reputation can invite more activism, firms that have suffered reputational shocks are further incentivized to protect themselves through self-regulation (Abito et al. 2016; Dorobantu, Henisz, and Narthey 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; McDonnell and King 2013; McDonnell et al. 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 buyers who have had their reputation compromised by past revelations about their suppliers' harmful practices are more likely than other buyers to make efforts to ensure that their suppliers correct abuses. We therefore hypothesize:

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

Design of Monitoring Program

A monitoring system has numerous components, from the stringency of the underlying substantive standards to the frequency and rigor of inspections to the composition of the inspection team. All of these design features have implications for suppliers' social compliance and compliance improvement. We focus on auditor training and the pre-announcement of audits because, as we argue below, these are features that might facilitate the transfer of compliance-related knowledge from auditors to suppliers, thus promoting improved labor practices.

Auditor Training

There is much skepticism about whether social auditing can foster improvement, and questions have been raised about the competence of 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's perceived interests (LeBaron and Lister 2015), and that some are outright corrupt (Clean Clothes Campaign 2005). Others believe that auditors are easily misled by suppliers that maintain fake wage and hours records and coach their workers to lie to auditors

about working conditions (Power, Ng, and Singh 2008; AFL-CIO 2013; LeBaron and Lister 2015). Some, including critics of social auditing, have suggested that more highly trained auditors could be more effective (Locke et al. 2009; AFL-CIO 2013), and research has indeed found that better-trained auditors identify more violations (Short et al. 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 discovered (Amengual 2010). Our conversations with auditors and managers at social auditing firms indicate that the auditors' training typically teaches them how to find violations and what conditions tend to cause them. Such training helps auditors identify root causes and develop compliance solutions. Recent evidence indicates 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. Studies in the knowledge transfer literature find that certain types of training can improve the ability to 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; Reinhold, Pedersen, and Foss 2011). We therefore hypothesize:

Hypothesis 3 (H3): Suppliers will improve their labor practices more following code-of-conduct audits conducted by audit teams that are more highly trained.

Signaling a Cooperative Approach to Social Auditing

Significant debate surrounds the approach buyers should take to audits, including whether they should be conducted in a policing style or a more cooperative style and, relatedly, whether they should be announced in advance. Some argue that pre-announcing gives suppliers time to cover up bad behavior (Clean Clothes Campaign 2005; AFL-CIO 2013; LeBaron and Lister 2015), and there is empirical evidence for that (Gray 2006; Marks 2012; Toffel et al. 2015). Worker rights advocates therefore long have favored unannounced audits (Frenkel and Scott 2002).

Although it seems clear that unannounced audits will reveal more information about supplier wrongdoing, it is less clear whether they will foster improvement. It is possible that suppliers might be motivated to improve their practices if they know that they can be caught at any time through a surprise audit. However, most buyers do not conduct audits regularly enough for unannounced audits to operate as a serious deterrent. Moreover, studies have suggested that a punitive, policing-style approach to monitoring can undermine compliance by dampening intrinsic motivations to comply (Short and Toffel 2010). That approach can also foster resentment among the regulated

community that can lead to backlash against regulatory requirements (Bardach and Kagan 1982; Ayres and Braithwaite 1992). Social auditors report that conducting unannounced audits “to catch managers unaware . . . aggravates the relationship between buyers and suppliers” and that such “tensions make it difficult to achieve any sustainable change” (Gould 2005: 28).

A consensus has begun to emerge among academics and practitioners that suppliers are more likely to improve with a less punitive, more cooperative approach to monitoring (Locke et al. 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 et al. 2009: 321). The underlying theory, developed most extensively by Ayres and Braithwaite in *Responsive Regulation* (1992), is that regulators’ signals of 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).

Although we do not observe the micro-level interactions among the buyers, suppliers, and auditors in our data and so cannot assess whether these parties have trusting or cooperative relationships, we argue that buyers formally signal trust and a cooperative approach to monitoring when they give suppliers advance notice of audits. Our conversations with ethical supply chain managers and social auditors 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. At the very least, announcing an audit indicates trust in the formal economic 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). To be clear, our argument is not that providing advance notice of audits will *cause* suppliers to improve more rapidly; instead, we argue that announced audits signal to suppliers that they are subject to a more cooperative (less punitive) monitoring regime. These suppliers are more likely than others to perceive that they are trusted by their buyers and, consequently, these suppliers will be more willing to reciprocate that perceived trust by investing in improvement. Thus, although an announced audit might uncover fewer violations, we hypothesize:

Hypothesis 4 (H4): Suppliers that receive advance notice of code-of-conduct audits will improve their labor practices more than other suppliers.

Auditor Training in the Context of a Cooperative Approach to Auditing

Auditing might be particularly effective in improving supplier practices when more knowledgeable auditors engage with suppliers that are 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 (Szulanski 1996; Dyer and Chu 2003; McEvily, Perrone, and Zaheer 2003). For instance, Cheng, Yeh, and Tu (2008) showed 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 argued earlier in the article, announcing audits signals trust in the supplier; we therefore expect suppliers who receive advance notice to be more receptive to the knowledge auditors 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 their labor practices more following announced code-of-conduct audits than following unannounced audits.

Recognizing that the hypothesized conditions might lead different types of violations to improve at disparate rates, we include below an extension to our analysis that disaggregates our dependent variable—improvement across all labor practices—to better understand which types of labor practices improve more (or less) rapidly under which conditions.

Data and Method

Empirical Context and Sample

We tested our hypotheses using data from code-of-conduct audits that were carried out by a large social auditing company (henceforth, the “social auditor”) that requested anonymity. During our sample period, the social auditor served Fortune 500 companies and was accredited to conduct audits of several leading social compliance standards. It operated in more than 100 countries and its staff spoke more than 20 languages. 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

Table 1. Industry Composition of Audits and Audited Suppliers

<i>Industry</i>	<i>Audits</i>		<i>Suppliers</i>	
	<i>Number</i>	<i>Percentage</i>	<i>Number</i>	<i>Percentage</i>
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

the differences between these codes are slight, which gave us confidence in treating all of them in the same manner.

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 include at least two audits. Our estimation sample consists of 8,677 focal audits conducted at 4,940 suppliers spanning 13 industries in 66 countries. In our sample, factories are audited an average of every 202 days, with an interquartile range of 83 to 293 days. Those categorized as annual audits are conducted every 344 days on average. The most common industries in our sample are garments, accessories, electronics, and toys (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) (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 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.

Dependent Variable

Our dependent variable measures the change in a supplier's compliance between its prior and focal audits. Audit data are the only available

²Nearly 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.

Table 2. Location of Audits and Audited Suppliers

Supplier country	Audits		Suppliers	
	Number	Percentage	Number	Percentage
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

large-scale measure of supplier compliance with private labor standards, and audits are the most commonly used measure of labor standards compliance in the literature on global value chains (e.g., Locke et al. 2007; Oka 2010a, 2010b; Ang et al. 2012; Distelhorst et al. 2015; Toffel et al. 2015).

To avoid undue influence of outliers, we used a metric akin to the difference in log violations but which is calculable even when violation counts are zero. To calculate *improvement*, we divide the number of violations from the focal audit plus 1 by the number of violations from the previous audit plus 1, take the natural log of that ratio, and then multiply 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 (there are 75 possible violations across these domains) and where $V_{i,t-1}$ is the comparable figure from that

supplier's prior audit.³ We add 1, the minimum non-zero value in the data, 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 percentage change metric, which can be overly sensitive to outliers and can inflate large changes.⁶ Our log form reduces skewness⁷ 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 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 in a given year, 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)—which

³The maximum possible number violations in each category were: child labor (7), forced or compulsory labor (5), working hours (7), occupational safety and health (31), minimum wage (15), treatment of foreign workers and subcontractors (4), and disciplinary practices (6). We excluded violation categories that, according to our data provider, do not apply to all suppliers (dormitory conditions and canteen violations) or were interpreted dissimilarly by auditors in different countries (freedom of association, the right to organize and bargain collectively, legal or client requirements), or where research has shown codes to diverge from one another in their details (freedom of association, nondiscrimination clauses) (O'Rourke 2002; Rodríguez-Garavito 2005). Our decision to omit violations of freedom of association and collective bargaining is also supported by research demonstrating that auditors often fail to properly identify these types of violations (Anner 2012; Egels-Zandén and Merk 2014).

⁴Though 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. Adding 0.1 to the numerator and denominator, instead of adding 1, yielded nearly identical results.

⁵For 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.

⁶For example, skewness declines by a factor of 10, from a value of 4.2 for percentage change to a mere 0.4 for our *improvement* metric, and kurtosis declines by a factor of nearly 7, from 30.5 for percentage change to 4.5 for *improvement*.

⁷The 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 and 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.

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. Both the press and NGOs are critical actors in the transnational advocacy networks that we theorize will generate exposure risks, thus it is important to include both variables in our analysis. Because press freedom and NGO density are highly correlated ($\rho = 0.83$), we use principle components analysis as a data reduction technique (Hair, Anderson, Tatham, and Black 1998; Kennedy 2008), 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% of the variance between press freedom and NGO density. We refer to this first component as *supplier environmental pressure*.⁸

Institutional pressure to improve labor standards is directed not only at suppliers but also—in fact, largely—at multinational buyers. To operationalize that dimension of institutional pressure, we rely on negative media reports, as others have done for similar purposes (Fiaschi, Giuliani, and Nieri 2013; Kölbel, Busch, and Jancso 2017). In particular, we 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). 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 subscribers, including activists, businesses, governments, global media, and investors. BHRRC invites companies to respond to any post that names them and reports that 86% of companies do so (Business & Human Rights Resource Centre 2017), suggesting that companies are attentive to these reports. We measure *targeted buyer pressure* by taking the number of times the buyer on whose behalf the audit was conducted appeared in articles in this database during the year prior to the audit, adding 1, and logging that sum to reduce skew.⁹

⁸A robustness test that includes *press freedom* and *NGO density* in our models instead of *supplier environmental pressure* yields the same inferences as our primary models.

⁹Because many buyers appeared in no such articles, we conducted robustness tests that measured this using a dummy (rather than a count) 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. Estimates yielded nearly identical results when we added this dummy to our primary specification and when we substituted it for our primary measure of *pressure on buyers (prior audit)*.

We calculate auditor training as the number of audit training courses an auditor had taken, based on data provided by the social auditor. Auditors are trained on various topics, including: audit skills to help auditors identify violations; substantive issues relevant to a specific industry, region, or supplier (for example, subcontracting in the garment industry); and the requirements of specific auditing protocols that certain clients have adopted, such as SA8000. Because audits are typically conducted by an audit team, we measure *maximum auditor training* as the largest number of training courses that any one 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.¹⁰ The maximum number of training courses for audit teams averaged 6.9. We use the maximum training of any one team member because this measures the greatest potential to identify code-of-conduct violations and to transfer knowledge on how to remediate them.

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 0 for unannounced audits, based on data from the social auditor. Whether an audit is announced or unannounced is typically determined by the buyer. In our sample, 76% of the audits were announced.¹¹

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 an opportunity set that differs from those with a “cleaner” history, which may influence their likelihood of improvement. Audits in our sample have a maximum of 75 violations, but such high counts were very rare. *Violations (prior audit)* is the number of 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 then logged (after adding 1).

¹⁰Robustness tests (not reported) indicate that using a team’s average auditor training instead of its maximum auditor training yields nearly identical results.

¹¹We found no evidence that a supplier’s prior violation count or the duration of the 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 duration 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.

Because prior research indicates that auditing is less stringent when suppliers pay their own auditors (Jiang, Stanford, and Xie 2012; Duflo, Greenstone, Pande, and Ryan 2013; Short and Toffel 2016), we created 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* (when we lacked information about who paid).

Re-audits typically have a more limited scope, tending to focus on concerns raised at the prior audit. Because this could mechanically affect improvement rates, we include three dummy variables as controls: *prior audit was re-audit, but focal audit was not*; *focal audit was re-audit, but prior audit was not*; *prior and focal audit were re-audits*. The baseline condition is that neither was a re-audit.

Our interviews with social auditors—at the firm that provided our 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.¹² We therefore control for *audit duration (prior audit)*, which we calculated by taking the log (after adding 1) of the number of staff hours required to conduct the prior audit.

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 count variable to denote each supplier's first audit in the estimation sample, its second audit, and so on.¹³ In our models, we flexibly control for *audit sequence* by including a dummy for each value, which avoids imposing the assumption that audit sequence 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*. We use dummies instead of a ratio because 82% of our sample's audit teams

¹²We 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).

¹³A supplier's first observation in the estimation sample (*audit sequence* = 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* = 1 for 57% of the observations in our estimation sample, 2 for 22%, 3 for 10%, and 4 or more for 10%.

were single gender: all-female (50%) and all-male (32%). The remainder were evenly split (15%), or had another mixed ratio (3%).

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).¹⁴

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 prior audit. We include only the prior audit values for these country-level variables because they are very stable over the period of time between two consecutive supplier audits. 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.¹⁵

Because domestic legal protections for labor rights could influence suppliers' perception of how much pressure they are under we obtain *labor laws* scores from Mosley (2011).¹⁶ These scores measure the extent to which domestic law provides collective labor rights such as the right 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 this index to measure the stringency of country-level workers' rights protections generally (e.g., Greenhill, Mosley, and Prakash 2009; Dean 2015; Toffel, Short, and Ouellet 2015; Fransen and Burgoon 2017) on the basis that collective rights

¹⁴Robustness tests (not reported) indicate that using teams' *average* auditor tenure yields nearly identical results.

¹⁵*FDI inflows* measures net inflows of foreign direct investment (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 companies 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.

¹⁶We 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.

are foundational to other workers' rights, for example, 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 creating *buyer employment (prior audit)*, logging employment to reduce skew. To do so, we obtained annual values of employment from Amadeus, Capital IQ, Hoovers, or Thomson ONE Banker.¹⁷

Summary statistics are reported in Table 3. Correlations are reported in Table A.1 in the Online Appendix.

Estimation and Results

We test our hypotheses by estimating a model that predicts *improvement* based on the independent and control variables described above and several additional control variables explained below. Our model uses log versions of our continuous independent and control variables (except *supplier environmental pressure*, which was created using principle components analysis) both to reflect our sense that they have a diminishing marginal influence on improvement and to diminish the potential impact of outliers. In our specification, each observation includes variables measured at an establishment's focal audit and prior audit; therefore, all establishments in our estimation sample have had at least two audits.

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 focal audit, which is used to construct our dependent variable. Since failing to account for these factors could bias our estimates, we also control for *maximum auditor training (focal audit)* and *announced (focal audit)*. 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 the prior one. Similarly, 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 advanced warning allowed it to fix or hide problems. We do not include the focal-audit values of *supplier environmental pressure* or *targeted buyer pressure* because they are very stable over time—their respective correlations between prior and focal audits is 0.99 and 0.89—and including them would substantially increase multicollinearity while adding almost no new information.

¹⁷The 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 data set.

Table 3. Summary Statistics

	<i>Mean</i>	<i>SD</i>	<i>Minimum</i>	<i>Maximum</i>
Improvement	0.22	0.86	-3.26	3.26
Supplier environmental pressure (prior audit)	-0.34	1.19	-1.14	4.02
Targeted buyer pressure (prior audit)	0.38	0.90	0	5.55
Maximum auditor training [§] (prior audit)	0	1.00	-1.65	1.64
Announced (prior audit)	0.76	0.43	0	1.00
Maximum auditor training [†] (focal audit)	2.04	0.87	0	3.47
Announced (focal audit)	0.77	0.42	0	1.00
Violations (prior audit)	7.23	5.81	0	25.00
Violations [†] (prior audit)	1.84	0.77	0	3.26
Paid by supplier or third party (prior audit)	0.51	0.50	0	1.00
Paid by the buyer (prior audit)	0.43	0.49	0	1.00
Paid by unknown entity (prior audit)	0.06	0.24	0	1.00
Paid by supplier or third party (focal audit)	0.52	0.5	0	1.00
Paid by the buyer (focal audit)	0.42	0.49	0	1.00
Paid by unknown entity (focal audit)	0.05	0.22	0	1.00
Prior audit was re-audit, but focal audit was not	0.10	0.30	0	1.00
Focal audit was re-audit, but prior audit was not	0.38	0.49	0	1.00
Prior and focal audit were re-audits	0.19	0.40	0	1.00
Audit duration [†] (prior audit)	1.44	0.43	0.03	4.39
Previous auditor (prior audit)	0.20	0.40	0	1.00
Previous auditor (focal audit)	0.23	0.42	0	1.00
Audit sequence	1.84	1.31	1.00	11.00
All-female audit team (prior audit)	0.53	0.50	0	1.00
Mixed-gender audit team (prior audit)	0.16	0.37	0	1.00
All-female audit team (focal audit)	0.50	0.50	0	1.00
Mixed-gender audit team (focal audit)	0.18	0.38	0	1.00
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
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
Buyer employment [†] (prior year)	10.32	1.22	1.61	14.35

Notes: † indicates logged. § indicates logged, then standardized. $N = 8,677$ audits, except 7,774 for targeted buyer pressure, 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. FDI, foreign direct investment; GDP, gross domestic product; SD, standard deviation.

Because several audit design elements and audit team characteristics at the prior audit could influence violations recorded in that audit, and because 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 the 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. Because suppliers might respond differently to buyers in institutional contexts exerting varying levels

of pressure (Toffel, Short, and Ouellet 2015), we include fixed effects for buyer countries.¹⁸ We log *maximum auditor training (prior audit)* and *audit duration (prior audit)* to reduce skew and then standardize them to facilitate an elasticity interpretation of coefficients in response to a one standard deviation change. We use the log form of all other continuous variables to facilitate their interpretations as elasticities.

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% improvement rate (calculated as $1.6 \div 7.2$) corresponds to the sample average *improvement* rate of 0.22 reported in the summary statistics (Table 3).¹⁹

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.

We test H1–H4 with model 1 and report results in Table 4. The statistically significant positive coefficient on *supplier environmental pressure (prior audit)* ($\beta = 0.089$; $p < 0.01$) indicates that suppliers tend to improve more in countries in which civil society monitoring has greater potential to expose noncompliance with codes of conduct, which supports H1. The coefficient magnitude indicates that a one-standard-deviation increase in *supplier environmental pressure (prior audit)* (such as a change from Vietnam to the Philippines) is associated with an increase in *improvement* from an average of 22% to 32.6%, based on average predictions across our sample.²⁰ This 32.6% 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 sample average reduction of 1.6. Support for H1 is robust to replacing *supplier environmental pressure (prior audit)* with its two underlying elements, which yields statistically significant positive coefficients on both standardized *press freedom (prior audit)* and standardized *NGO density (prior audit)*. This provides further evidence 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.

¹⁸While 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 for all other buyer country attributes that are relatively stable during our sample period.

¹⁹Of the variation in the number of *violations*, 16.3% comes from the factory-country level, 48.3% comes from the factory level, and the remaining 35.4% comes from other factors at the observation level. We also decomposed variation in *improvement* (our dependent variable) by using a mixed model to estimate our primary specification and found that 5.2% of the variation in the number of violations comes from the factory-country level, 5.0% comes from the factory level, and the remaining 89.9% comes from other factors at the observation level.

²⁰The 32.6% is calculated by adding to 0.22 the product of 0.089 (the coefficient on *supplier environmental pressure*) and 1.19 (the standard deviation of *supplier environmental pressure*).

Table 4. Regression Results

<i>Dependent variable</i>	(1)	(2)
	<i>Improvement</i>	
Supplier environmental pressure (prior audit)	0.089** [0.015]	0.088** [0.015]
Targeted buyer pressure (prior audit)	0.037** [0.010]	0.039** [0.010]
Maximum auditor training ^s (prior audit)	0.031** [0.012]	-0.008 [0.018]
Announced (prior audit)	0.049** [0.011]	0.049** [0.013]
Maximum auditor training ^s (prior audit) × Announced (prior audit)		0.051** [0.011]
Maximum auditor training [†] (focal audit)	-0.053** [0.016]	-0.054** [0.016]
Announced (focal audit)	0.048** [0.017]	0.046* [0.018]
Violations [†] (prior audit)	0.676** [0.025]	0.679** [0.026]
Paid by supplier or third party (prior audit)	-0.016 [0.021]	-0.020 [0.020]
Paid by supplier or third party (focal audit)	0.136** [0.017]	0.141** [0.016]
Paid by unknown entity (prior audit)	0.009 [0.022]	0.008 [0.022]
Paid by unknown entity (focal audit)	-0.027 [0.065]	-0.025 [0.063]
Prior audit was re-audit, but focal audit was not	-0.121** [0.035]	-0.124** [0.033]
Focal audit was re-audit, but prior audit was not	0.277** [0.016]	0.274** [0.016]
Prior and focal audit were re-audits	0.197** [0.011]	0.194** [0.012]
Audit duration [†] (prior audit)	0.120** [0.027]	0.126** [0.026]
Previous auditor (prior audit)	0.013 [0.026]	0.014 [0.026]
Previous auditor (focal audit)	0.032+ [0.018]	0.033+ [0.018]
All-female audit team (prior audit)	0.009 [0.013]	0.009 [0.013]
Mixed-gender audit team (prior audit)	-0.019 [0.025]	-0.019 [0.025]
All-female audit team (focal audit)	-0.063** [0.018]	-0.063** [0.018]
Mixed-gender audit team (focal audit)	-0.122** [0.027]	-0.122** [0.027]
Maximum auditor tenure [†] (prior audit)	-0.195 [0.195]	-0.188 [0.192]
Maximum auditor tenure [†] (prior audit), squared	0.051 [0.051]	0.048 [0.050]
Maximum auditor tenure [†] (focal audit)	-0.574** [0.171]	-0.584** [0.171]
Maximum auditor tenure [†] (focal audit), squared	0.154** [0.043]	0.157** [0.043]

(continued)

Table 4. Continued

Dependent variable	(1)	(2)
	Improvement	
FDI inflows [†] (prior audit)	-0.064 [0.040]	-0.063 [0.040]
Labor laws [†] (2002)	0.157 [0.115]	0.158 [0.116]
GDP per capita [†] (prior audit)	0.051** [0.016]	0.053** [0.015]
GDP per capita in buyer country [†] (prior audit)	-0.007 [0.030]	0.004 [0.031]
Buyer employment [†] (prior year)	-0.043** [0.005]	-0.045** [0.005]
Year fixed effects	Included	Included
Sequence fixed effects	Included	Included
Industry fixed effects	Included	Included
Buyer-country fixed effects	Included	Included
R-squared	0.41	0.42

Notes: Ordinary least squares (OLS) regression coefficients with standard errors clustered by supplier country in brackets. $N = 8,677$ observations (each based on two consecutive audits) from 4,940 factories. † indicates logged. § indicates logged, then standardized. Baseline (omitted) categories are *paid by the buyer* for focal and prior audit when neither was a re-audit. All models include dummy variables to indicate instances in which the following variables were missing data and thus recoded to 0: *targeted buyer pressure (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). FDI, foreign direct investment; GDP, gross domestic product.

** $p < 0.01$; * $p < 0.05$; + $p < 0.10$.

The statistically significant positive coefficient on *targeted buyer pressure (prior audit)* ($\beta = 0.037$; $p < 0.01$) illustrates greater average improvement for suppliers to buyers that have already been publicly exposed for harms to workers in their supply chain, which supports H2. The coefficient magnitude indicates that a one-standard-deviation increase in *targeted buyer pressure (prior audit)* is associated with an increase in *improvement* from an average of 22% to 23.2%, based on average predictions across our sample.²¹

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 H3.²² The coefficient magnitude 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

²¹The 23.2% is calculated by adding to 0.22 the product of 0.031 (the coefficient on *supplier environmental pressure*) and 0.40 (its standard deviation).

²²Our finding that a better-trained audit team 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.

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% 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.

The statistically significant positive coefficient on *announced (prior audit)* ($\beta = 0.049$; $p < 0.01$) indicates that greater improvement follows announced audits than unannounced audits, which supports H4.²³ Predictive margins indicate that suppliers whose prior audit was announced experienced an average 23.2% improvement, compared to 18.2% for suppliers whose prior audit was unannounced. Applied to the average 7.23 violations in the prior audit, this is an average decline of 1.67 violations following announced audits versus 1.34 after unannounced audits. This average differential of 0.33 violations per audit corresponds to one more violation being mitigated after three announced audits than after three unannounced audits. These results are largely identified based on differences between factories because only a small fraction of the factories in our sample have prior audits that are a mix of announced and unannounced audits.

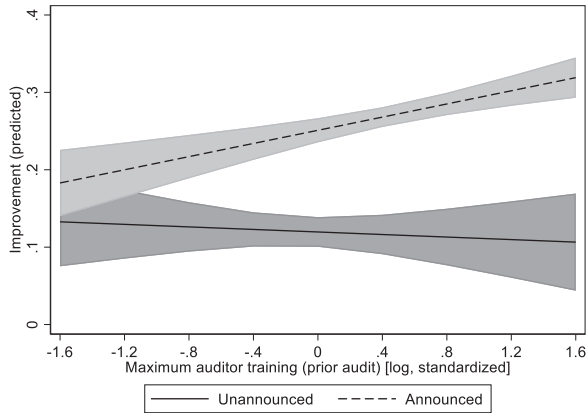
To test H5, we add a term that interacts *maximum auditor training (prior audit)* and *announced (prior audit)* and report the results as model 2 in Table 4. The statistically significant positive coefficient on the interaction term ($\beta = 0.051$; $p < 0.01$) indicates that better-trained audit teams 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 in which 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.

Supplementary Analysis

We conducted several additional analyses to assess the robustness of our results. Our results are robust to several alternative estimation approaches, including: two-way clustering standard errors at the supplier-country and buyer-company levels; a cross-sectional model (i.e., one observation per factory) using factory means of all variables; and a random-effects model using factory-level random effects (not reported). While we believe our

²³If announcing the prior audit gave factories time to hide or solve problems, 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.

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



Notes: 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 with all other variables at their actual values. Shading depicts 95% confidence intervals.

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 from our primary specifications. These negative binomial regression²⁴ results, reported in Table A.2 in the Online Appendix, confirm all inferences from our primary models and thereby indicate that our results are robust to this alternative specification.²⁵

Because improvement might depend on the time between the prior and focal audits, we re-estimated our models predicting *improvement rate*, an alternative dependent variable that explicitly accounted for the time lag since the prior audit. *Improvement rate* is calculated by dividing *improvement* (our primary dependent variable) by the log of the number of days since the prior audit. These results, too, yield statistically significant support for

²⁴Because the dependent variable of this model, number of violations, exhibits overdispersion—the ratio of the variance to the mean is 4.4 (that is, 24.7 / 5.6)—we followed the conventional approach of using negative binomial regression rather than Poisson. Moreover, a likelihood-ratio test yields a chi-squared value of 6,931, which indicates that the probability that we would observe these data conditional on alpha equaling 0 (an assumption underlying the Poisson estimator) is virtually zero.

²⁵We also estimated OLS models that predicted two other alternative dependent variables: the difference in violations between the focal and prior audit and the difference between the logs of these values. These models yielded results substantially similar to those of our primary models and of the alternative models reported in Table A.2 in the Online Appendix. Specifically, they yielded the same inferences for our institutional features and stronger evidence of auditor training, but only evidence of *announced* having an effect when it was interacted with *auditor training*.

our hypotheses except that auditor training remains a significant predictor of improvement only in the presence of announced audits.

We coded our variables irrespective of whether the focal and prior audit were conducted for the same brand. Estimating our primary models on the subsample of 6,458 observations in which an establishment's focal and prior audits were conducted for the same brand yielded insights similar to those of our primary approach. The one difference is that the estimates on this subsample yield evidence of an auditor training effect only in the presence of an announced audit (H5), but not an overall effect (H3).

Because so many of the observations in our sample are from China, we examined whether our primary results held when estimating our models on just the 6,294 observations from the 3,378 factories in China. To do so, we omit our factory-country variables and the country-year-level variables (to avoid multicollinearity concerns, as they changed only slightly over time) and cluster standard errors by factory. All the hypothesized effects continue to yield statistically significant coefficients of the same sign and similar magnitude (though in some cases with larger standard errors, likely attributable in part to the smaller sample size).

We estimated models that accounted for the fact that some buyers always sought unannounced audits, some always sought announced audits, and some sought a mix. Specifically, we added two control variables to our primary models: one dummy variable indicating audits on behalf of buyers that always specified announced inspections and another indicating audits on behalf of 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. Overall, our primary results proved markedly stable throughout various robustness tests.

Extensions

Our primary analysis measures improvement based on the difference between the total number of violations cited in the focal audit and in the prior audit, aggregating several types of violation to capture improvement comprehensively. Recognizing that different types of violations might improve at different rates under the hypothesized conditions (Barrientos and Smith 2007; Ruwanpura 2012; Egels-Zandén and Lindholm 2015; Stroehle 2017), 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 of the violation categories making up our *improvement* variable for which at least 10% of audits exhibited variation; these were child labor, working hours, minimum wage, and occupational safety and health

(OSH).²⁶ We created an improvement metric for each category by applying the same formula used to create our primary improvement metric. The correlation between these four variables ranges from 0.20 to 0.34. 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 and also indicate that our hypothesized variables are associated with varying degrees of improvement across different violation types, thus highlighting important nuances of our primary results.

First, we find that under conditions of greater institutional pressure in suppliers' domestic environments, significantly more improvement occurs in all four subcategories (child labor, working hours, minimum wages, and OSH violations). This finding confirms the importance of institutional pressure to improvement on all of these dimensions.

Second, we find more improvement in working hours, minimum wage, and OSH violations among suppliers of buyers directly targeted by institutional pressure. These findings support our intuition that reputation-sensitive buyers will be more proactive in seeking improvements to supply chain working conditions. OSH violations have been a focus of activist organizing and media exposés, making them highly salient to brands with reputation concerns; these brands may apply more pressure to improve supplier conditions. Working hours and wages are dimensions that buyer sourcing practices can influence, and buyers facing more institutional pressure might be especially attentive to this. For example, such buyers might be more prone to avoiding frequent change orders or delaying orders, thus reducing suppliers' need to work excessive hours. Such practices can also mitigate the risk of workplace injuries. Similarly, buyers concerned about their reputations might be more likely to avoid bargaining to minimize suppliers' profit margins to the point that suppliers feel compelled to constrain wages in ways that violate codes of conduct. We were somewhat surprised that we did not find evidence of greater improvements in high-reputation-salience violations such as child labor among suppliers of buyers facing more institutional pressure. Nonetheless, we note that the coefficient on this violation type remains positive and its lack of significance may be due to statistical power: Factories in our sample were much less likely to exhibit variation over time in the number of child labor violations than in the other types of violations. Of the observations in our sample, 66% have the same number of child labor violations in the focal and prior audits, compared to 17 to 33% for violations of working hours, minimum wage, and OSH.

²⁶The scores for *disciplinary practices*, *forced or compulsory labor*, and *treatment of foreign workers and subcontractors* changed in fewer than 10% of audits. To avoid generalizing from such limited variation, we did not estimate models to predict those three categories.

Table 5. Extension Results

<i>Dependent variable</i>	(1)	(2)	(3)	(4)
	<i>Child labor score improvement</i>	<i>Working hours score improvement</i>	<i>Minimum wage score improvement</i>	<i>OSH score improvement</i>
Supplier environmental pressure (prior audit)	0.044** [0.006]	0.034** [0.012]	0.060** [0.009]	0.043* [0.017]
Targeted buyer pressure (prior audit)	0.007 [0.006]	0.021** [0.007]	0.017** [0.005]	0.033** [0.008]
Maximum auditor training ^s (prior audit)	0.017* [0.008]	0.010 [0.009]	0.034** [0.006]	0.026* [0.012]
Announced (prior audit)	-0.013** [0.004]	0.011 [0.010]	0.010+ [0.006]	0.041** [0.011]
Maximum auditor training [†] (focal audit)	-0.004 [0.005]	-0.045** [0.010]	-0.015+ [0.008]	-0.044+ [0.024]
Announced (focal audit)	0.036** [0.012]	0.014 [0.016]	0.053** [0.013]	0.043** [0.015]
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)	-0.001 [0.006]	0.024** [0.007]	-0.043+ [0.022]	-0.034* [0.015]
Paid by supplier or third party (focal audit)	0.028** [0.007]	0.048** [0.015]	0.081** [0.023]	0.034+ [0.018]
Paid by unknown entity (prior audit)	0.001 [0.006]	0.011 [0.027]	0.004 [0.016]	0.049* [0.024]
Paid by unknown entity (focal audit)	0.006 [0.010]	-0.023 [0.037]	-0.018 [0.039]	-0.021 [0.052]
Prior audit was re-audit, but focal audit was not	-0.010 [0.011]	-0.050** [0.013]	-0.055** [0.010]	-0.191** [0.054]
Focal audit was re-audit, but prior audit was not	0.029** [0.008]	0.028** [0.009]	0.072** [0.013]	0.304** [0.024]
Prior and focal audit were re-audits	0.037** [0.005]	0.006 [0.009]	0.041** [0.013]	0.243** [0.010]
Audit duration [†] (prior audit)	0.049** [0.016]	0.088** [0.017]	0.084** [0.015]	0.160** [0.040]
Previous auditor (prior audit)	0.007 [0.005]	-0.000 [0.017]	-0.012 [0.017]	0.020 [0.026]
Previous auditor (focal audit)	-0.011 [0.013]	-0.014 [0.013]	0.017 [0.022]	0.037* [0.017]
All-female audit team (prior audit)	0.021** [0.007]	0.002 [0.010]	-0.009 [0.013]	0.007 [0.012]
Mixed-gender audit team (prior audit)	0.018* [0.008]	-0.011 [0.012]	-0.025 [0.020]	0.007 [0.035]
All-female audit team (focal audit)	-0.004 [0.009]	-0.035* [0.015]	-0.065** [0.011]	-0.044* [0.020]
Mixed-gender audit team (focal audit)	0.011 [0.014]	-0.051* [0.020]	-0.067** [0.018]	-0.127** [0.030]
Maximum auditor tenure [†] (prior audit)	0.048 [0.064]	0.120 [0.095]	0.057 [0.096]	0.272 [0.336]

(continued)

Table 5. Continued

<i>Dependent variable</i>	(1)	(2)	(3)	(4)
	<i>Child labor score improvement</i>	<i>Working hours score improvement</i>	<i>Minimum wage score improvement</i>	<i>OSH score improvement</i>
Maximum auditor tenure [†] (prior audit), squared	-0.006 [0.017]	-0.037 [0.028]	-0.012 [0.029]	-0.071 [0.092]
Maximum auditor tenure [†] (focal audit)	-0.175** [0.050]	-0.269* [0.114]	-0.145 [0.101]	-0.412* [0.172]
Maximum auditor tenure [†] (focal audit), squared	0.043** [0.014]	0.062* [0.030]	0.037 [0.031]	0.125** [0.045]
FDI inflows [†] (prior audit)	-0.026* [0.013]	-0.067* [0.027]	-0.030 [0.022]	-0.022 [0.042]
Labor laws [†] (2002)	0.010 [0.046]	0.183 [0.122]	0.020 [0.071]	-0.006 [0.102]
GDP per capita [†] (prior audit)	0.000 [0.006]	0.044** [0.014]	0.035** [0.010]	0.029 [0.017]
GDP per capita in buyer country [†] (prior audit)	0.014 [0.009]	-0.042 [0.026]	0.031 [0.019]	0.022 [0.020]
Buyer employment [†] (prior year)	-0.009** [0.003]	-0.023** [0.005]	-0.004 [0.004]	-0.025** [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
<i>R</i> -squared	0.44	0.36	0.32	0.31
Mean <i>y</i> in sample	0.072	0.055	0.081	0.246

Notes: Ordinary least squares (OLS) regression coefficients with standard errors clustered by supplier country in brackets. $N = 8,335$ observations (each based on two consecutive audits) from 4,870 factories, slightly smaller than in Table 4 because of a few missing values of violation category counts. † indicates logged. § indicates logged, then standardized. See Table 4 for additional notes. FDI, foreign direct investment; GDP, gross domestic product; OSH, occupational safety and health.

** $p < 0.01$; * $p < 0.05$; + $p < 0.10$.

Third, we find that visits by highly trained auditors lead to significantly more improvement in child labor, wages, and OSH scores. These findings are consistent with our hypothesis that better-trained auditors can better convey compliance information. Improvements in child labor scores may be attributable to suppliers' greater willingness to receive and follow advice from more highly trained auditors. 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.

Finally, we find that supplier improvement following announced audits is driven primarily by improvement in OSH violations—precisely the kind 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 are grounds for contract termination, an exchange of information is typically not required to convey compliance expectations and practices. We therefore do not expect formal 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 of trust.

Discussion

Our findings reveal several important structural conditions under which codes of conduct and monitoring regimes adopted by TNCs are associated with improvements in suppliers' labor practices. 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 discovered and publicized; the more institutional pressure targeted at buyers from additional revelations, the more their suppliers improve. In addition, we find that suppliers improve more not only with external institutional pressures but also when monitoring programs are designed in ways that facilitate knowledge transfer. Suppliers improved to a greater degree following audits with advance notice, particularly in areas such as OSH, where compliance assistance can be most helpful. We also find suppliers are more likely to improve when their auditors are highly trained, but only when those auditors conduct pre-announced audits. These findings make important contributions to the literatures on supply chain labor standards regulation, private politics, and decoupling.

First, 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; LeBaron and Lister 2015), suggesting the important role auditors can play given the proper tools. While some have suggested that better-trained auditors can be more effective (Locke, Amengual, and Mangla 2009; AFL-CIO 2013), to our knowledge ours is the first study to empirically document this conjecture in this context. At the same time, our finding that even highly trained auditors add no significant compliance improvement through unannounced audits suggests the limitations of training and the need to consider program design holistically rather than piecemeal.

Second, our finding that pre-announced audits were followed by greater improvement in OSH practices but not in child labor practices adds nuance to the debate surrounding whether audits should be announced or unannounced. The improvement in OSH practices is consistent with 1) the prediction that compliance can develop iteratively in response to cooperative gestures by those implementing the rules (e.g., Axelrod 1984; Scholz 1984; Ayres and Braithwaite 1992) and 2) 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 that is 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; LeBaron and Lister 2015; Short, Toffel, and Hugill 2016).

These findings add an important empirical dimension to this highly polarized and largely theoretical debate by documenting the difficult sets of tradeoffs involved in designing monitoring regimes. There can be a tradeoff between violation discovery and compliance improvement. Thus, buyers must consider which to prioritize depending on the aims of their monitoring regimes, such as whether they primarily seek to collect the most complete and accurate information, to catch suppliers committing particularly harmful violations, or to improve working conditions by creating conditions for cooperation with suppliers. In addition, there is a risk that the signals buyers send to foster cooperation and improvement might be perceived by some suppliers as laxity, encouraging noncompliance. Buyers who seek to foster cooperation and improvement should think carefully about the signals they send to suppliers to ensure that those signals are not misinterpreted as license to violate the code. Our finding that highly trained auditors were associated with accelerated improvement in child labor practices as well as in other violation categories suggests that these tradeoffs might be reconciled to some degree through better training of auditors. Further research should explore how monitoring approaches can be deployed and combined to leverage their comparative advantages.

Third, the extension disaggregating our dependent variable contributes to a body of research addressing how improvement varies across different violation categories. Studies have shown, for instance, that labor code compliance tends to improve more rapidly in some categories, such as health and safety, than in others, such as freedom of association (Barrientos and Smith 2007; Ruwanpura 2012; Egels-Zandén and Lindholm 2015; Stroehle 2017). Our findings extend existing research that identifies conditions associated with improvement in wages and hours requirements but not OSH requirements. Oka (2015) found that unionized suppliers improve their compliance with wage, hours, and leave standards much more substantially than with OSH standards because unions in developing economies—where most suppliers grappling with code compliance are located—tend to prioritize pocketbook issues over OSH issues. Distelhorst et al. (2017) similarly found that wages and hours practices, but not OSH compliance, improve in factories that adopt lean manufacturing practices. We extend this work by identifying institutional and program design features associated with improvements in OSH violations as well as these other categories. These

findings suggest the importance of alternative pathways of influence to improve working conditions overall.

Fourth, our finding of greater improvement among suppliers to reputation-compromised buyers suggests that exposure by activists can have substantive impacts beyond the largely symbolic responses identified in the private politics literature. Prior research documents that activism prompts firms to adopt symbolic structures such as “impression management tactics” (McDonnell and King 2013: 411), public “concessions” to conform to activists’ demands (Eesley and Lenox 2006; King 2008), and CSR officer positions or board committees (McDonnell et al. 2015). However, this research has not revealed whether activism is related to changes in organizational behavior that align it more closely with the activists’ normative goals. Our finding that buyers tainted by negative media coverage are 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. Moreover, our finding that the *more* pressure buyers experience in their institutional environment, the *more* likely their suppliers are to improve is an important strategic insight for activists considering how to target their resources. Research has shown that firms that are repeatedly targeted by activists become more receptive to future activist challenges, in part through the adoption of formal corporate structures to manage social responsibility issues (McDonnell et al. 2015). We demonstrate that frequent targets may go beyond formal organizational responses to actually improve the practices that prompted activists’ objections.

Finally, demonstrating the relationship between program design features and performance improvement adds an important dimension to understandings of decoupling. 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 design features create contingencies for one another. 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 premise by suggesting that the difference between symbol and substance may depend, in part, on how organizational structures are designed.

Limitations and Future Research

Our study has limitations but also invites promising future research. 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 interventions, such as more stringent government regulation, legally binding international standards, or labor union activities. These are vital research questions.

Our findings are subject to several data limitations. Violations cited in social audits are an imperfect measure of objective labor conditions. Audits present an incomplete snapshot of factory practices, and research suggests that social auditing suffers from biases due to the misaligned incentives of auditors, inadequate training of auditors, and suppliers' elaborate efforts at subterfuge (O'Rourke 2002; Esbenshade 2004; LeBaron and Lister 2015). In our analysis, we make extensive efforts to address potential biases by controlling for several factors known to affect auditors' detection of violations, including who pays for the audit, ongoing auditor-supplier relationships, and the gender composition of audit teams (Short, Toffel, and Hugill 2016) and by omitting from our data those violation categories such as freedom of association and collective bargaining where auditors have been found notoriously incapable of reliably identifying violations (Anner 2012; Egels-Zandén and Merk 2014).

Moreover, we examine factories that faced at least two social audits by a single firm. By omitting those 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 it does not enable us to compare practices between auditing firms. Omitting certain types of code violation from our analysis enhanced the reliability of our improvement measure, but it 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 as firm size and regulatory enforcement practices. Although we believe that our proxies for key independent variables are reasonable, we cannot rule out the possibility that they are incomplete. Finally, it is possible that alternative governance structures might influence improvement, including whether audits are conducted by brand staff or by a third-party auditing firm and whether codes are sponsored by brands, multi-stakeholder regimes, or NGOs. Our study paves the way for others to examine these and other factors.

Conclusion

As the anti-sweatshop movement makes the "TNC into the central locus of struggle over labor rights and globalization" (Bartley and Child 2014: 657), it is crucial to understand whether formal organizational structures such as codes of conduct and supplier monitoring can produce meaningful social change. We identify conditions at the institutional and program design levels under which these formal organizational structures are associated with measurable improvements in working conditions. Our findings suggest

key considerations that should inform social monitoring and activist targeting strategies aimed at raising labor standards in global supply chains.

References

- Abito, Jose M., David Besanko, and Daniel Diermeier. 2016. Corporate reputational dynamics, private regulation, and activist pressure. *Advances in Strategic Management* 34: 235–99.
- AFL-CIO. 2013. Responsibility outsourced: Social audits, workplace certification and twenty years of failure to protect. Accessed at <http://www.aflcio.org/Learn-About-Unions/Global-Labor-Movement/Responsibility-Outsourced-Report> (December 18, 2014). Washington, DC: The American Federation of Labor and Congress of Industrial Organizations.
- Amengual, Matthew. 2010. Complementary labor regulation: The uncoordinated combination of state and private regulators in the Dominican Republic. *World Development* 38(3): 405–14.
- Amengual, Matthew, and Laura Chirot. 2016. Reinforcing the state: Transnational and state labor regulation in Indonesia. *ILR Review* 69(5): 1056–80.
- Ang, Debra, Drusilla Brown, Rajeev Dehejia, and Raymond Robertson. 2012. Public disclosure, reputation sensitivity, and labor law compliance: Evidence from Better Factories Cambodia. *Review of Development Economics* 16(4): 594–607.
- Anner, Mark. 2012. Corporate social responsibility and freedom of association rights: The precarious quest for legitimacy and control in global supply chains. *Politics & Society* 40(4): 609–44.
- Axelrod, Robert. 1984. *The Evolution of Cooperation*. New York: Basic.
- Ayres, Ian, and John Braithwaite. 1992. *Responsive Regulation: Transcending the Deregulation Debate*. New York: Oxford University Press.
- Bardach, Eugene, and Robert A. Kagan. 1982 [2002]. *Going by the Book: The Problem of Regulatory Unreasonableness*. New Brunswick, NJ: Transaction.
- Barkemeyer, Ralf. 2009. Beyond compliance—below expectations? CSR in the context of international development. *Business Ethics: A European Review* 18(3): 273–89.
- Barrientos, Stephanie, and Sally Smith. 2007. Do workers benefit from ethical trade? Assessing codes of labour practice in global production systems. *Third World Quarterly* 28(4): 713–29.
- Bartley, Tim, and Curtis Child. 2014. Shaming the corporation: The social production of targets and the anti-sweatshop movement. *American Sociological Review* 79(4): 653–79.
- Bartley, Tim, and Niklas 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.
- . 2016. Beyond decoupling: Unions and the leveraging of corporate social responsibility in Indonesia. *Socio-Economic Review* 2: 231–255.
- Berliner, Daniel, Anne R. Greenleaf, Milli Lake, Margaret Levi, and Jennifer 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.
- Bird, Yanhua, Jodi L. Short, and Michael W. Toffel. 2019. Coupling labor codes of conduct and supplier labor practices: The role of internal structural conditions. *Organization Science* 30(4): 647–867.
- Boiral, Olivier. 2007. Corporate greening through ISO 14001: A rational myth? *Organizational Science* 18(1): 127–46.
- Borgatti, Stephen P., and Rob Cross. 2003. A relational view of information seeking and learning in social networks. *Management Science* 49(4): 432–45.
- Bromley, Patricia, and Walter 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. Accessed at <https://www.business-humanrights.org/en/about-us> (August 19, 2017). London, UK.

- Cannizzaro, Anthony P., and Robert J. Weiner. 2015. Multinational investment and voluntary disclosure: Project-level evidence from the petroleum industry. *Accounting, Organizations and Society* 42: 32–47.
- Chatterji, Aaron K., and Michael W. Toffel. 2010. How firms respond to being rated. *Strategic Management Journal* 31(9): 917–45.
- Cheng, Jaohong H., Chunghsing H. Yeh, and Chiawen W. Tu. 2008. Trust and knowledge sharing in green supply chains. *Supply Chain Management: An International Journal* 13(4): 283–95.
- Chih, Hsiang L., Hsiang H. Chih, and Tzu Y. Chen. 2010. On the determinants of corporate social responsibility: International evidence on the financial industry. *Journal of Business Ethics* 93(1): 115–35.
- Clean Clothes Campaign. 2005. Looking for a quick fix: How weak social auditing is keeping workers in sweatshops. Accessed at <https://www.evb.ch/fileadmin/files/documents/Konsum/Quickfix05.pdf> (May 31, 2016). Amsterdam, The Netherlands.
- Dean, Adam. 2015. Power over profits: The political economy of workers and wages. *Politics & Society* 43(3): 333–60.
- Distelhorst, Greg, Jens Hainmueller, and Richard M. Locke. 2017. Does lean performance improve labor standards? Management and social performance in the Nike supply chain. *Management Science* 63(3): 707–28.
- Distelhorst, Greg, Richard M. Locke, Timea Pal, and Hiram Samel. 2015. Production goes global, compliance stays local: Private regulation in the global electronics industry. *Regulation & Governance* 9(3): 224–42.
- Dobbin, Frank, and Erin L. Kelly. 2007. How to stop harassment: Professional construction of legal compliance in organizations. *American Journal of Sociology* 112(4): 1203–43.
- Dorobantu, Sinziana, Witold J. Henisz, and Lite Nartey. 2017. Not all sparks light a fire: Stakeholder and shareholder reactions to critical events in contested markets. *Administrative Science Quarterly* 62(3): 561–97.
- Duflo, Esther, Michael Greenstone, Rohini Pande, and Nicholas Ryan. 2013. Truth-telling by third-party auditors and the response of polluting firms: Experimental evidence from India. *Quarterly Journal of Economics* 128(4): 1499–545.
- Dyer, Jeff H., and Wujin 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.
- Eesley, Charles, and Michael J. Lenox. 2006. Firm responses to secondary stakeholder action. *Strategic Management Journal* 27(8): 765–81.
- Egels-Zandén, Niklas. 2007. Suppliers' compliance with MNCs' codes of conduct: Behind the scenes at Chinese toy suppliers. *Journal of Business Ethics* 75(1): 45–62.
- Egels-Zandén, Niklas, and Henrik Lindholm. 2015. Do codes of conduct improve worker rights in supply chains? A study of Fair Wear Foundation. *Journal of Cleaner Production* 107(16): 31–40.
- Egels-Zandén, Niklas, and Jeroen Merk. 2014. Private regulation and trade union rights: Why codes of conduct have limited impact on trade union rights. *Journal of Business Ethics* 123(3): 461–73.
- Elevate. 2016. ELEVATE. Accessed at <http://www.elevatelimited.com/> (March 15, 2016). Hong Kong, China.
- Esbenshade, Jill. 2004. *Monitoring Sweatshops: Workers, Consumers, and the Global Apparel Industry*. Philadelphia: Temple University Press.
- Espeland, Wendy N. 1998. *The Struggle for Water: Politics, Rationality, and Identity in the American Southwest*. Chicago: University of Chicago Press.
- Faccio, Mara. 2006. Politically connected firms. *American Economic Review* 96(1): 369–86.
- Fiaschi, Davide, Elisa Giuliani, and Federica 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, Luc. 2012. *Corporate Social Responsibility and Global Labor Standards: Firms and Activists in the Making of Private Regulation*. New York: Routledge.

- Fransen, Luc, and Brian Burgoon. 2017. Support for ethical consumerism and welfare states in the global economy: Complements or substitutes? *Global Policy* 8(S3): 42–55.
- Frenkel, Stephen J., and Duncan Scott. 2002. Compliance, collaboration, and codes of labor practice: The Adidas connection. *California Management Review* 45(1): 29–49.
- Frynas, Jędrzej George. 2005. The false developmental promise of Corporate Social Responsibility: Evidence from multinational oil companies. *International Affairs* 81(3): 581–98.
- Gereffi, Gary, John Humphrey, and Timothy Sturgeon. 2005. The governance of global value chains. *Review of International Political Economy* 12(1): 78–104.
- Gould, Daniella. 2005. The problem with supplier audits. *Corporate Responsibility Management* 2(1): 24–29.
- Gray, Garry C. 2006. The regulation of corporate violations: Punishment, compliance, and the blurring of responsibility. *British Journal of Criminology* 46(5): 875–92.
- Greene, William H. 2007. *Econometric Analysis*, 6th edition. Upper Saddle River, NJ: Prentice Hall.
- Greenhill, Brian, Layna Mosley, and Aseem Prakash. 2009. Trade-based diffusion of labor rights: A panel study, 1986–2002. *American Political Science Review* 103(4): 169–90.
- Guillén, Mauro F., and Laurence Capron. 2015. State capacity, minority shareholder protections, and stock market development. *Administrative Science Quarterly* 61(1): 125–60.
- Gulati, Ranjay, and Maxim 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, Emilie M., and Kiyoteru Tsutsui. 2005. Human rights in a globalizing world: The paradox of empty promises. *American Journal of Sociology* 110(5): 1373–411.
- Hair, Joseph F., Jr., Rolph E. Anderson, Ronald L. Tatham, and William C. Black. 1998. *Multivariate Data Analysis*, 5th edition. Delhi: Pearson Education.
- Hallett, Tim. 2010. The myth incarnate: Recoupling processes, turmoil, and inhabited institutions in an urban elementary school. *American Sociological Review* 75(1): 52–74.
- Intertek. 2016. Global supplier management system. Accessed at <http://www.intertek.com/consumer/auditing/global-supplier-management/> (May 31, 2016).
- Jiang, John (X.), Mary H. Stanford, and Yuan Xie. 2012. Does it matter who pays for bond ratings? Historical evidence. *Journal of Finance Economics* 105(3): 607–21.
- Justice, Derefaka W. 2006. Corporate social responsibility and society's expectations of business. *Perspectives on Work* 9(2): 14–16.
- Kalev, Alexandra, Frank Dobbin, and Erin Kelly. 2006. Best practices or best guesses? Assessing the efficacy of corporate affirmative action and diversity policies. *American Sociological Review* 71(4): 589–617.
- Keck, Margaret E., and Kathryn Sikkink. 1998. *Activists beyond Borders*. Ithaca, NY: Cornell University Press.
- Kennedy, Peter. 2008. *A Guide to Econometrics*, 6th edition. Malden, MA: Wiley-Blackwell.
- King, Brayden G. 2008. A political mediation model of corporate response to social movement activism. *Administrative Science Quarterly* 53(3): 395–421.
- . 2014. Reputational dynamics of private regulation. *Socio-Economic Review* 12(1): 200–206.
- Kölbl, Julian F., Timo Busch, and Leonhardt M. Jancso. 2017. How media coverage of corporate social irresponsibility increases financial risk. *Strategic Management Journal* 38(11): 2266–84.
- Kotchen, Matthew, and Jon J. Moon. 2012. Corporate social responsibility for irresponsibility. *B.E. Journal of Economic Analysis & Policy (Contributions)* 12(1): Article 55.
- LeBaron, Genevieve, and Jane Lister. 2015. Benchmarking global supply chains: The power of the “ethical audit” regime. *Review of International Studies* 41(special issue 5): 905–24.
- Levine, David I., Michael W. Toffel, and Matthew S. Johnson. 2012. Randomized government safety inspections reduce worker injuries with no detectable job loss. *Science* 336(6083): 907–11.
- Lim, Alwyn, and Kiyoteru 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, Richard M. 2013. *The Promise and Limits of Private Power: Promoting Labor Standards in a Global Economy*. Cambridge, UK: Cambridge University Press.
- Locke, Richard, Matthew Amengual, and Akshay Mangla. 2009. Virtue out of necessity? Compliance, commitment, and the improvement of labor conditions in global supply chains. *Politics & Society* 37(3): 319–51.
- Locke, Richard M., Fei Qin, and Alberto Brause. 2007. Does monitoring improve labor standards? Lessons from Nike. *Industrial and Labor Relations Review* 61(1): 3–31.
- Locke, Richard M., Ben A. Rissing, and Timea 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): 519–52.
- Locke, Richard M., and Monica Romis. 2007. Improving work conditions in a global supply chain. *MIT Sloan Management Review* 48(2): 54–62.
- Loewenstein, Jeffrey, Leigh Thompson, and Dedre Gentner. 2003. Analogical learning in negotiation teams: Comparing cases promotes learning and transfer. *Academy of Management Learning and Education* 2(2): 119–27.
- Luo, Jiao, Stephan Meier, and Felix Oberholzer-Gee. 2012. No news is good news: CSR strategy and newspaper coverage of negative firm events. Harvard Business School Working Paper 12-091. Boston, MA: Harvard Business School.
- Marks, Kathy. 2012. Factory workers are “forced to lie” during Adidas safety inspections. *Independent* (UK) (April 13). Accessed at <http://www.independent.co.uk/news/world/asia/factory-workers-are-forced-to-lie-during-adidas-safety-inspections-7644018.html> (May 4, 2016).
- Marquis, Christopher, and Cuili Qian. 2014. Corporate social responsibility reporting in China: Symbol or substance? *Organization Science* 25(1): 127–48.
- Mayer, Roger C., James H. Davis, and F. David Schoorman. 1995. An integrative model of organizational trust. *Academy of Management Review* 20(3): 709–34.
- McBarnet, Doreen. 2007. Corporate social responsibility beyond law, through law, for law: The new corporate accountability. In Doreen McBarnet, Aurora Voiculescu, and Tom Campbell (Eds.), *The New Corporate Accountability: Corporate Social Responsibility and the Law*, pp. 9–56. Cambridge, UK: Cambridge University Press.
- McDonnell, Mary-Hunter, and Brayden G. King. 2013. Keeping up appearances: Reputation threat and prosocial responses to social movement boycotts. *Administrative Science Quarterly* 58(3): 387–419.
- McDonnell, Mary-Hunter, Brayden G. King, and Sarah A. Soule. 2015. A dynamic process model of private politics: Activist targeting and corporate receptivity to social challenges. *American Sociological Review* 80(3): 654–78.
- McEvily, Bill, Vincenzo Perrone, and Akbar Zaheer. 2003. Trust as an organizing principle. *Organization Science* 14(1): 91–103.
- Meyer, John W., John Boli, George M. Thomas, and Francisco O. Ramirez. 1997. World society and the nation-state. *American Journal of Sociology* 103(1): 144–81.
- Meyer, John W., and Brian Rowan. 1977. Institutionalized organizations: Formal structure as myth and ceremony. *American Journal of Sociology* 83(2): 340–63.
- Mosley, Layna. 2011. Replication data for collective labor rights dataset. Accessed at <http://hdl.handle.net/1902.1/15590> (August 29, 2013). Harvard Dataverse.
- Nadler, Janice, Leigh Thompson, and Leaf Van Boven. 2003. Learning negotiation skills: Four models of knowledge creation and transfer. *Management Science* 49(4): 529–40.
- Nadvi, Khalid, Peter Lund-Thomsen, Hong Xue, and Navjote Khara. 2011. Playing against China: Global value chains and labour standards in the international sports goods industry. *Global Networks* 11(3): 334–54.
- Oka, Chikako. 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.
- . 2010b. Channels of buyer influence and labor standard compliance: The case of Cambodia’s garment sector. In David Lewin, Bruce E. Kaufman, and Paul J. Gollan (Eds.), *Advances in Industrial and Labor Relations*, vol. 17, pp. 153–83. Bingley, UK: Emerald Group.

- . 2015. Improving working conditions in garment supply chains: The role of unions in Cambodia. *British Journal of Industrial Relations* 54(3): 647–72.
- O'Rourke, Dara. 2002. Monitoring the monitors: A critique of corporate third-party labour monitoring. In Rhys Jenkins, Ruth Pearson, and Gill Seyfang (Eds.), *Corporate Responsibility and Labour Rights: Codes of Conduct in the Global Economy*, pp. 196–207. London: Earthscan.
- Overdevest, Christine. 2010. Comparing forest certification schemes: The case of ratcheting standards in the forest sector. *Socio-Economic Review* 8(1): 47–76.
- Perkins, Susan E. 2014. When does prior experience pay? Institutional experience and the multinational corporation. *Administrative Science Quarterly* 59(1): 145–81.
- Potoski, Matthew, and Aseem Prakash. 2005. Green clubs and voluntary governance: ISO 14001 and firms' regulatory compliance. *American Journal of Political Science* 49(2): 235–48.
- Power, Carla, I.-C. Ng, and M. Singh. 2008. The burden of good intentions. *Time International* 171(25): 45–48.
- Reinholt, Mia, Torben Pedersen, and Nicolai 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): 1277–97.
- Rodríguez-Garavito, Cesar. 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): 203–33.
- Ruggie, John. 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. Accessed at <http://www.reports-and-materials.org/Ruggie-report-7-Apr-2008.pdf> (July 5, 2015). Geneva: United Nations, Human Rights Council.
- Ruwanpura, Kanchana N. 2012. Ethical codes: Reality and rhetoric. A study of Sri Lanka's apparel sector. University of Southampton Working Paper. Accessed at <http://eprints.soton.ac.uk/337113/> (May 31, 2016).
- Scholz, John T. 1984. Cooperation, deterrence, and the ecology of regulatory enforcement. *Law & Society Review* 18(2): 179–224.
- Seidman, Gay. 2007. *Beyond the Boycott: Labor Rights, Human Rights, and Transnational Activism*. New York: Russell Sage Foundation.
- Shamir, Ronen. 2011. Socially responsible private regulation: World-culture or world-capitalism? *Law & Society Review* 45(2): 313–36.
- Shea, Anna, Mariko Nakayama, and Jody 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, Jodi L., and Michael W. Toffel. 2010. Making self-regulation more than merely symbolic: The critical role of the legal environment. *Administrative Science Quarterly* 55(3): 361–96.
- . 2016. The integrity of private third-party compliance monitoring. *Administrative & Regulatory Law News* 42(1): 22–25.
- Short, Jodi L., Michael W. Toffel, and Andrea R. Hugill. 2016. Monitoring global supply chains. *Strategic Management Journal* 37(9): 1878–97.
- Stroehle, Judith C. 2017. The enforcement of diverse labour standards through private governance: An assessment. *European Review of Labour and Research* 23(4): 475–93.
- Szulanski, Gabriel. 1996. Exploring internal stickiness: Impediments to the transfer of best practice within the firm. *Strategic Management Journal* 17(S2): 27–43.
- Thomas-Hunt, Melissa C., Tonya Y. Ogden, and Margaret A. Neale. 2003. Who's really sharing? Effects of social and expert status on knowledge exchange within groups. *Management Science* 49(4): 464–77.
- Thompson, Leigh, Dedre Gentner, and Jeffrey 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, Andras. 2010. From ritual to reality: Demography, ideology, and decoupling in a post-Communist government agency. *Academy of Management Journal* 53(6): 1474–98.

- Toffel, Michael W., Jodi L. Short, and Melissa Ouellet. 2015. Codes in context: How states, markets, and civil society shape adherence to global labor standards. *Regulation & Governance* 9(3): 205–23.
- UL Responsible Sourcing. 2015. Supply chain certification. Accessed at <http://services.ul.com/service/supply-chaincertification/> (May 15, 2016).
- Utting, Peter. 2005. Corporate responsibility and the movement of business. *Development in Practice* 15(3–4): 375–88.
- Weil, David, and Carlos 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.
- Zajak, Sabrina. 2017. *Transnational Activism, Global Labor Governance, and China*. New York: Palgrave Macmillan.