Second- versus Third-party Audit Quality: Evidence from Global Supply Chain Monitoring[†]

Maria R. Ibanez
Northwestern Kellogg School of Management
Evanston, IL 60208
maria.ibanez@northwestern.edu

Ashley Palmarozzo
Harvard Business School
Boston, MA 02163
apalmarozzo@hbs.edu

Jodi L. Short
UC Law, San Francisco
200 McAllister
San Francisco, CA
shortj@uchastings.edu

Michael W. Toffel
Harvard Business School
Boston, MA 02163
mtoffel@hbs.edu

August 7, 2024

Capitalizing on the superior credibility and flexibility and potential lower cost of external assessments, many global buyers are relying less on their own employee ("second-party") auditors and more on third-party auditors to monitor and prevent environmental and social misconduct in supply chains. Despite ingrained assumptions that third-party auditors' greater independence reduces bias and improves audit quality, there are concerns that this trend risks eroding audit quality. Drawing on agency theory for a more nuanced understanding of auditor incentives and on data from a global fashion brand, we find third-party auditors indeed less effective, especially as a given factory's region exhibits more reported corruption or less potential oversight by second-party auditors. Global buyers can bolster third-party–audit quality by increasing the presence of second-party auditors in a given region, emphasizing such deployment in more corrupt regions, and rotating amongst third-party audit firms. Our findings can inform better-designed monitoring not only of suppliers, but also of other business partners that create risks for brands, such as franchisees, distributors, vendors, and purchasing agents.

Keywords: supply chain management; auditing; working conditions; sustainability; empirical operations management

1. Introduction

Branded global buyers face increasing scrutiny and pressure from activists, shareholders, and consumers to avoid harmful working conditions at their suppliers' factories. Fashion brands, for example, have been condemned for dangerous or exploitive working conditions in their global supply chains—from fires and building collapses in Bangladesh to child labor in India to forced labor in the Xinjiang region of China. Because media exposés and activist mobilization can inflict significant reputational damage, global buyers routinely (a) require suppliers to adopt codes of labor conduct and (b) have their suppliers periodically audited for adherence to these codes. It is not clear, however, who should do the auditing.

There is concern among labor advocates, nongovernmental organizations (NGOs), and other external stakeholders about the quality of global buyers' auditing (e.g., Spicer 2021); specifically, that auditors will suppress information about harmful working conditions so that the global buyers and

[†] We thank the company, which opted to remain anonymous, that shared access to its supplier code-of-conduct audit data. We gratefully acknowledge the research assistance of Melissa Ouellet and helpful feedback from Xiang Ao, Hong Luo, Feng Zhu, and participants in the Alliance for Research on Corporate Sustainability (ARCS) Seminar, Duke Fuqua School of Business Strategy Seminar, and University of Texas at Austin McCombs School of Business Operations Management Seminar.

noncompliant suppliers who hire them can maintain the perceived competitive advantage of noncompliant production. This has led stakeholders to insist that auditors be independent of both the brands they serve and the suppliers they monitor and therefore to consider *third-party* auditors "the gold standard" (Prakash and Potoski 2007: 790). Third-party auditors are employed by audit firms whose primary business is auditing; having no employment relationship with the entities they audit or with the clients who hire them to do the auditing, they do not have the same incentives these companies might have to hide wrongdoing. This distinguishes them both from *first-party* auditors, employed by the audited entity (in our context, a supplier) to conduct self-audits, and from *second-party* (or "in-house") auditors, typically employed by a business partner (such as a global buyer) of the audited entity.[1] Third-party auditors are "widely acknowledged to be more competent, credible and transparent than a buyer's own audit" (Huq et al. 2016), leading stakeholders to demand that global buyers rely on such auditors to monitor suppliers' compliance with labor codes of conduct (e.g., McAllister 2012, Lebaron and Lister 2015).

In fact, global buyers have, in recent years, relied less on their own (second-party) auditors and more on third-party auditors. For example, Inditex (owner of fashion retailer Zara) increased the share of third-party audits from 63% in 2017 to 91% in 2021, conducting 64% fewer audits using its own employees despite increasing total audits by 40%.[2] During the same period, Nike raised the share of third-party audits of suppliers from 4% to 88%.[3] This shift not only satisfies stakeholder demands, but also enables global buyers to benefit from the greater flexibility and potentially lower cost of outsourcing audits to third parties. However, our conversations with several brands suggest that some are concerned about the quality of third-party audits and believe that their own (second-party) auditors can do better. We hypothesize why this might be so and potential ways to mitigate the quality gap.

Although third-party auditors can be biased by conflicts of interest arising from economic incentives to cultivate client loyalty through leniency (Duflo et al. 2013, Pierce and Toffel 2013, Toffel et al. 2015, Short et al. 2016)—especially in markets with high competition (Bennett et al. 2013)—little consideration has been given to second-party auditors as an alternative. In fact, the literature on social compliance tends to assume that in-house auditors will be even more biased than third-party auditors, being less independent. We complicate this assumption by examining how independence itself creates incentives that can undermine auditor performance. Specifically, we draw on insights from agency theory to theorize conditions under which auditor performance might improve under greater brand control.

A large fashion brand which, since it chooses to remain anonymous, we will refer to as

[1] We use the term *second-party* rather than *in-house* auditors to make clear that they are employees of the global buyer, not of the factories being audited.
[2] Authors' calculations based on data from Inditex's annual reports (Inditex 2017: 241; Inditex 2021: 512).
[3] Authors' calculations based on data from Nike's sustainability reports (Nike 2018: 43; Nike 2022: 150).

FashionCo, granted us access to more than ten thousand audits of thousands of its supply chain factories around the world over an 11-year period. Using this data, we find justification for the aforementioned concerns about third-party auditors and evidence of several factors that can mitigate these concerns. The audits monitor suppliers' compliance with FashionCo's Supplier Code of Conduct, which is closely aligned with the Ethical Trading Initiative's base code of labor practices and is broadly similar to codes used by other large global buyers. The code covers, for instance, wages-and-hours requirements, occupational health and safety standards, and protections against child labor, forced labor, and discrimination. Code violations at supplier factories are pervasive and detecting and reporting them requires effort on the part of auditors, who must cite evidence to support their findings. We therefore refer to audits that cite more violations at a given establishment, which more accurately represents actual workplace conditions and reflects more effort and less willingness to conceal violations—all else equal—as *higher-quality audits*.

Exploiting quasi-random assignment of audits to auditors—after conditioning on observable factors—and comparing audits conducted within the same supplier factories, we find that, on average, second-party auditors report 9.7% more violations than third-party auditors do. This calls into question the commonly held view that independent third-party auditors are better at monitoring suppliers. Two other findings reinforce the value of second-party audits. First, we find that increases in reported corruption within a region compromise third-party audits more than they do second-party audits. This makes second-party auditors especially valuable in protecting global brands from factory problems in regions where corruption is a greater concern. Second, we find that third-party auditors perform higher-quality audits in regions where second-party auditors conduct a greater proportion of the annual factory audits. This supports our hypothesis that a greater presence of second-party auditors disciplines third-party auditors by providing a credible threat of monitoring and backsourcing (that is, shifting from third parties to second parties). This finding (a) suggests that the trend toward reducing second-party audits risks lowering the quality of third-party audits and (b) implies that global buyers should maintain a sufficient presence of second-party auditors in regions that pose the greatest reputational risk. We find that rotating across various third-party audit firms for a given factory also bolsters third-party-audit quality. Overall, these results reveal risks of global buyers' shift from second- to third-party audits and several factors that can exacerbate—or ameliorate—such risks.

To our knowledge, this is the first study using large-scale field data to compare second-party and third-party social auditors. Our findings contribute to streams of literature on supply chain social responsibility, quality management and monitoring, the effect of bias on audit quality, and behavioral operations. Because social auditing is the operational lynchpin of many companies' social responsibility programs, our findings on determinants of its quality can help companies improve social responsibility

performance. Our results also contribute to research on auditing as a tool for process quality management. In addition, we contribute to literature on behavioral operations by identifying behavioral drivers of auditor performance that challenge conventional assumptions about auditor incentives grounded in research on auditor bias. From a practical perspective, our findings reveal novel ways to improve auditor performance and can help managers implement more effective strategies not only for their supply chains but also for other business partners such as franchisees, distributors, vendors, and purchasing agents. Moreover, our results challenge the notion that third-party auditors are necessarily superior and suggest that well-intended calls for global buyers to rely on third-party auditors might have unintended consequences. The trend of global brands increasingly relying on third-party audits might in fact reduce audit quality and raise—rather than lower—risks for workers at global supply chain factories.

2. Hypotheses

Global buyers can monitor their supply chains using second-party auditors (their own employees) or third-party auditors (from audit firms). Despite the two types using a common audit protocol and standard procedures to make their audits comparable, they face different incentives that could affect audit quality—the number of violations they detect and cite at a given supplier factory. Supplementing theories of auditor independence with insights from agency theory, we hypothesize the relative quality of second- and third-party audits and theorize how three factors affect it: (a) regional corruption, (b) regional density of second-party relative to third-party auditors, and (c) audit firm rotation.

2.1. Audit Quality of Second-party versus Third-party Auditors

Operations management research has examined how global brands can foster more responsible and sustainable supplier practices (e.g., Kraft et al. 2018, Buell and Kalkanci 2021), mostly with respect to environmental performance (Lo et al. 2018, Muthulingam et al. 2022) and labor conditions (Caro et al. 2021, Villena et al. 2021, Huang et al. 2022, Vanpoucke and Klassen forthcoming). Monitoring suppliers to promote responsible production is difficult because it targets difficult-to-observe practices and processes rather than more readily observable product attributes (e.g., Parmigiani 2007, Jia 2018). Research on how to improve the auditing process has yielded insights based on empirical analyses (e.g., Pierce and Snyder 2008, Schneider 2012, Bennett et al. 2013, Duflo et al. 2013, Pierce and Toffel 2013, Short et al. 2016, Ibanez and Toffel 2020) and analytical derivation (e.g., Plambeck and Taylor 2016, Chen et al. 2020, Huang et al. 2022). However, the empirical evidence has focused on *either* internal (second-party) or external (third-party) auditors or else been agnostic about this distinction. Little is therefore known about their relative effectiveness, the focus of our study.

Auditor independence has long been the touchstone of audit quality across a wide range of monitoring contexts (Levitt 2000, Ammenberg et al. 2001, McAllister 2012). The formal independence of

third-party auditors from clients and audited parties is thought to attenuate biases that might arise from conflicts of interest (Kraakman 1986, Kouakou et al. 2013, Lebaron and Lister 2015), giving third-party auditors more credibility (Huq et al. 2016). The theory is that independent, third-party auditors have less incentive to overlook a supplier's violations because their employer gains no direct financial benefit or competitive advantage from those violations (Coffee 2004). The emphasis on independence to incentivize high-quality auditing assumes that the core threat to audit quality is auditors covering up wrongdoing they find. This is a real threat in supply chain auditing—and in other contexts, such as financial auditing. However, lack of independence is not the only threat to audit quality.

Audit quality is likewise threatened when auditors cut corners and exert lower effort in order to reduce their costs, just as service quality suffers in any setting when workers or contractors cut corners (Oliva and Sterman 2001). Whether auditors are covering for clients or just cutting corners, the outcome is the same: they will fail to see or fail to report critical information and thus produce incomplete, misleading audit reports. While independence might mitigate the incentive to cover for clients, it does not mitigate the incentive to cut corners. Rather, independence might encourage it by making it easier to hide. In other words, although independence can reduce conflicts of interest that might prompt auditors to engage in cover-ups, it creates agency problems that can incentivize corner-cutting. Agency problems arise when one party (the principal) delegates work to another party (the agent) and information asymmetry prevents the principal from observing or accurately judging how well the agent conducts the work. Because the incentives of principal and agent are seldom fully aligned, a moral hazard problem often arises whereby the agent is motivated to shirk duties (Jensen and Meckling 1976, Klein et al. 1978).

We draw on agency theory to hypothesize how auditor independence creates incentives that might undermine audit quality in certain circumstances. Specifically, we contrast the notions of independence and control that ground justifications for third- versus second-party auditing, respectively. Conceptualizing branded global buyers as principals and auditors as agents, we argue that agency problems are particularly likely to lead to shirking in the social auditing context and that these problems intensify in third-party relationships, where there is greater distance between principal and agent.

Agency problems are exacerbated when it is more difficult for principals to monitor agents' performance—especially when the quality of the products (Parmigiani 2007) or services (Jia 2018) supplied by the agent is not readily apparent. Monitoring the quality of supply chain auditors' performance is such a case because auditing occurs in factories worldwide, often thousands of miles from the buyer's headquarters, and is based on the auditor's real-time observations and judgments, which cannot easily be replicated and checked by the global buyer. To mitigate this kind of problem, firms tend to insource functions that are difficult to observe because it is easier to monitor and incentivize their own employees than outside agents. For example, firms are more likely to assign sales and distribution

functions to their own employees, rather than outsource them, when sales practices or measurement challenges make performance monitoring more difficult (Anderson and Schmittlein 1984, John and Weitz 1988).

Agency problems can be particularly severe when an agent's poor performance creates greater reputational risk for the principal than for the agent. Such asymmetrical reputation risk exacerbates the misalignment of incentives because it prevents agents from fully internalizing the costs of their own poor performance. Again, to remedy this incentive misalignment, firms tend to insource functions they ordinarily would outsource when poor-quality work by an agent would expose them to high reputational risk that an agent is unlikely to fully internalize (Nickerson and Silverman 2003, Mayer 2006). Nickerson and Silverman (2003), for example, find that trucking companies tend to use their own employee-drivers rather than third-party drivers on high-stakes routes for which poor delivery performance seriously imperils not merely a single route but the company's reputation or its broader operations.

In our context, global brands face significantly more risk of reputational harm from exposés of suppliers' poor labor practices than do third-party auditors who failed to catch or report them. The press and activists are more likely to target high-profile brands for their relationships with abusive suppliers, as this garners more attention and is more likely to provoke an industry response than exposing lesser-known audit firms (Bartley and Child 2014, King and McDonnell 2015). As providers of business-to-business services, firms providing third-party audit services are not household names and are therefore unlikely to be tagged by the press or activists, even if their shoddy work failed to uncover poor conditions that led to worker harm.[4] By contrast, second-party auditors steeped in a global buyer company's culture might develop a sense of loyalty to their own company that third-party auditors are unlikely to develop toward the many clients for whom they work only temporarily. Such loyalty to the buyer-company that employs them might prompt second-party auditors to better internalize the risks their company faces.

Third, reputational discipline is especially unlikely to mitigate agency problems between principals and third-party agents in the social auditing context. There is some (albeit mixed) evidence of reputational discipline in markets for financial auditing, where third-party financial auditors must publicly attest to the validity of their reports and are required to issue public restatements of their mistakes (Francis and Wilson 1988, Ljungqvist et al. 2007). However, it is infeasible for global buyers to impose reputational discipline on third-party social auditors because the work of social auditors is almost always

[4] Examples of typical media reports describing poor working conditions in global supply chains that blame factory owners and brands but do not (or only scarcely) mention auditors include articles about worker suicides at Apple supplier Foxconn (Merchant 2017), the fatal Tazreen factory fire (Manik and Yardley 2012), and the fatal Rana Plaza factory collapse (Rushe 2016). Similarly, NGOs focusing on global supply chain working conditions—such as the Worker Rights Consortium and the Fair Labor Association—only occasionally express skepticism about auditors and are far more likely to name brands than audit companies.

confidential and global buyers do not routinely publicly reveal the identity of their auditors. There is no systematic way for global buyers to learn which third-party social auditors make more mistakes or have been dismissed by other buyers. Thus, even if third-party social auditors have some incentive to provide higher-quality auditing to protect their branded clients from reputational risk, they do not face sufficient reputational pressure themselves to fully internalize this risk.

In addition to these agency problems, there are practical reasons why second-party auditors could be more effective than third-party auditors. First, whereas third-party auditors apply a variety of codes of conduct used by different global buyers and industries, second-party auditors apply only their employer's code of conduct, and such specialization can lead to better performance (e.g., Narayanan et al. 2009). Second, second-party auditors might receive greater respect and cooperation from audited suppliers because suppliers (a) might perceive them to have more authority and sway with the branded buyer and (b) might have the opportunity to develop bilateral trust over time (Huq et al. 2016). Cooperation is particularly important in the social auditing context because comprehensive assessment requires access to documents, worker interviewees, and all areas of the factory.

We therefore hypothesize:

Hypothesis 1 (H1): Second-party auditors will exhibit higher audit quality than third-party auditors.

2.2. Audit Quality of Third-party Auditors Is More Vulnerable to Regional Corruption

Many global supply chain factories are in regions of rampant corruption and bribery. Noncompliant suppliers may attempt to bribe auditors to pass audits (Chen et al. 2020); some factory managers in China have reported that "it is very common to offer bribes or other services to auditors" (China Labor Watch 2009). Suppliers in more corrupt areas could perceive concealing information and misleading auditors to be less unethical—or not unethical at all—making it harder for auditors to detect violations (Plambeck and Taylor 2016). For example, clothing suppliers in Bangladesh, a country known for high levels of corruption, often conceal nonconformance by keeping falsified duplicate records of hours worked, reporting this behavior as a clever, valuable skill of their compliance managers (Huq et al. 2016). Similarly, managers of several factories in China meeting with the founder and executive director of the NGO China Labor Watch described falsifying audits as a managerial necessity without ethical connotations; they complained that not falsifying audits was raising their costs and making it difficult to compete with factories that could pass audits through bribery and deception (China Labor Watch 2009). For all these reasons, audits of factories in more corrupt environments may yield fewer violations.

We integrate insights from our prior arguments and the literature on the corruption risk of auditing to argue that second- and third-party auditors differ in their vulnerability to corruption. We argue that increases in reported corruption in a region are less likely to compromise the effectiveness of second-party than of third-party auditors. Second-party auditors are more directly controlled and monitored by

the branded company that employs them and thus less subject to agency problems. They are also more likely to be loyal to the global brand that employs them and to share its incentives to avoid reputational harm from scandals that could arise from poor quality audits and corruption. Moreover, a supplier may be less inclined to try to bribe its buyer's own auditors, fearing that the buyer is more likely to find out and end the relationship.

We therefore hypothesize:

Hypothesis 2 (H2): Increases in reported corruption in a region will be associated with a sharper decrease in the quality of audits conducted by third-party auditors than by global buyers' second-party auditors.

2.3. Audit Quality of Third-party Auditors Improves in the Shadow of Second-party Auditors

Some buyers that extensively audit their supply chains use a mix of second- and third-party auditors. We propose that second-party auditors can be deployed to raise overall auditing quality because third-party auditors will detect and record more violations when there are more second-party auditors in the region. First, second-party auditors provide local competition for third-party auditors, which can deter shirking and improve performance. The more second-party auditors in a given market, the more plausible the competitive threat that the buyer can use them rather than an unsatisfactory third-party auditor. Firms commonly use dual-sourcing strategies—for instance, sourcing from multiple external suppliers—to generate competition and bolster supplier performance (e.g., Klotz and Chatterjee 1995, Yang et al. 2012).

Second, more second-party auditors in a region facilitates the brand's ability to monitor its third-party auditors in that region—just as franchisors face lower costs for monitoring physically proximate franchisees (Rubin 1978, Norton 1988). For instance, second-party auditors can spot-check third-party auditors' work by visiting the same factories. Even if extensive spot-checking is impractical, the possibility of being spot-checked could deter third-party auditors from shirking and thus improve their audits.

Using second-party auditors also helps buyers learn about the auditing process and better evaluate third-party auditors' performance. It has been theorized, for example, that sourcing from internal and external producers can help manufacturers evaluate the performance of third-party suppliers by comparing it to that of internal suppliers. Firms can tap second-party production knowledge to better assess whether defective products from a third-party supplier are due "to genuine manufacturing problems or to surreptitious supplier cheating" (Heide et al. 2014: 1166). Such sourcing strategies also enable manufacturers to benchmark performance between second-party and third-party suppliers so as to ratchet up performance by creating "a virtuous cycle of continuous improvement" (Puranam et al. 2013: 1151). Similarly, in franchising, company-owned units' performance can be used as a benchmark to motivate

franchisee-owned units to improve (e.g., Bradach 1997). We posit that, in the supply chain auditing context, more second-party auditing teams in a region will similarly motivate improvement by third-party auditors.

Based on these competitive and monitoring pressures, we hypothesize:

Hypothesis 3 (H3): Increases in the prevalence of second-party audits in a region will be associated with an increase in the quality of audits conducted by third-party auditors in that region.

2.4. Third-party Audit Firm Rotation Improves Audit Quality

Greater use of second-party auditors could also, however, increase the costs of auditing and decrease the efficiencies that often motivate outsourcing this function in the first place. Here, we argue that global buyers can use audit scheduling practices to bolster third-party-audit quality. Specifically, we hypothesize that rotating among third-party audit *firms*—so that the firm conducting a factory's next audit is not the one that conducted the previous one—will result in more violations cited.

Research suggests that ongoing relationships between *individual* auditors and the firms they repeatedly audit can create biases that undermine audit quality (Lennox et al. 2014, Short et al. 2016, Ball et al. 2017, Jin and Lee 2018). Some regulators therefore mandate periodic rotation of individual auditors (e.g., US Securities and Exchange Commission 2003), a practice commonly referred to in auditing scholarship and policy as "auditor rotation."

But would *firm* rotation offer benefits above and beyond those of rotating individual auditors? In an extensive review of the auditor rotation literature, Lennox (2014: 103) concludes that "we still do not have a clear idea as to whether mandatory audit firm rotation would make audit quality better or worse" and that, beyond the effects of rotating individual lead auditors, it remains unknown "whether audit firms *also* should be periodically changed." Many studies framed in terms of audit firm rotation do not, in fact, study that but instead examine the duration of audit firm relationships with the entities they audit (e.g., Carcello and Nagy 2004, Al-Thuneibat et al. 2011, Tepalagul and Lin 2015).

We propose that rotating audit firms will improve detection of violations above and beyond the benefits already associated with rotating individual auditors from the same firm. An auditor's performance is affected not only by *individual* emotional and cognitive factors but also by the *organizational* structures of their firm. Specifically, a company's culture, routines, and other organizational structures are likely to shape its employees' cognitive heuristics via mechanisms such as organizational norms (Trevino et al. 1998), codes of conduct (Cressey and Moore 1983), training (Metzger et al. 1993), and social networks (Brass et al. 1998).

We therefore argue that auditors are likely to detect more violations when neither they nor their coworkers conducted the entity's prior audit. For example, auditors shaped by their firm's organizational structures might, in consecutive factory visits, focus on similar violation types, overly search the same

areas of the factory previously covered by their colleagues (where previously discovered issues are more likely to have already been resolved), and under-search areas that might have revealed violations that were not stressed in their company training. Varying the firms auditing a factory over time should address these issues because different companies are likely to have different organizational structures, lessening the danger of cognitive limitations consistently constraining auditors in consecutive visits.

We therefore hypothesize:

Hypothesis 4 (H4): When consecutive audits of a given establishment are conducted by third-party auditors, the second audit's quality will be higher when rotating audit firms than when rotating only lead auditors from the same audit firm.

3. Data and Empirical Setting

3.1. Empirical Context and Measures

We obtained data from a large global fashion brand (via a data-sharing agreement) on all its global supply chain factory audits from 2007 to 2017. We refer to the company as FashionCo because it opted to remain anonymous. It operates hundreds of stores across the United States and Europe, with billions of dollars of global sales. It buys from supply chain factories in dozens of countries, requiring all to adhere to its code of conduct and be subjected to annual audits to assess their conformity. The code of conduct contains roughly 90 binary compliance elements spanning 12 violation categories: child labor, communication of the code of conduct to workers, discrimination, harsh worker treatment, environmental requirements, freedom of association, legal requirements, living wage, regular employment, working conditions, working hours, and voluntary employment. These categories track the Ethical Trading Initiative's Base Code, itself based on International Labour Organization core conventions. After auditors assess a factory's compliance, FashionCo assigns it an overall "social audit rating"—Red-Critical, Red, Amber, or Green—based on the number and severity of the cited violations.

Each audit is conducted either by FashionCo's own employees (second-party auditors) or by third-party auditors. All auditors use the same audit protocol and process: they review documents, tour the factory to observe conditions, interview workers, and record violations.[5] All auditors have access to the factory's prior audit records through an online system. The company either assigns a second-party auditor team or selects a third-party auditing firm (as described below) and pays for all audits. The vast majority of both second- and third-party auditors are locals hired from and located in the factory's country or region.[6] Ninety-five percent of third-party audits are conducted by for-profit audit firms and the rest by

[5] The number of auditor-days an audit requires is determined by the factory's size (measured by number of workers excluding management), in accordance with Sedex Members Ethical Trade Audit (SMETA) methodology.
[6] FashionCo selects third-party auditors based on their capabilities, capacity, geographic coverage, skills profiles, and other factors.

NGOs, intergovernmental organizations, or firms whose name in the dataset was blank. FashionCo indicated that these third-party audit firms are industry leaders and conduct audits for multiple global buyers.

Our unit of analysis is a factory audit. Because our model uses lags of several independent variables (i.e., information from the prior audit, which is missing for a factory's first audit), our estimation sample excludes the 15,775 audits in the raw data that were each factory's first audit in our dataset, which left 11,099 audits. Because we use factory–fixed-effect specifications (described below), the 947 audits of factories that had only one audit in our remaining data also drop out of our sample, leaving 10,152 observations. We dropped the 55 observations for audits that were conducted in the seven countries for which we both lacked World Bank bribery data (described below) and had fewer than 20 audits, leaving 10,097 observations. Finally, our Poisson model dropped the 19 observations pertaining to factories that always had zero violations. Our estimation sample thus includes 10,078 audits of 2,313 factories in 169 regions in 34 countries during 2007–2017. Second-party auditors conducted 27% and third-party auditors conducted the other 73%. Most of the supplier factories were in China (49% of the audits), India (18%), and Bangladesh (10%). FashionCo does not own any of the supplier factories and nearly all concurrently supplied other buyers.

3.2. Dependent Variable

Our dependent variable is the number of code-of-conduct *violations cited* in each audit report. Our measure is thus similar to those in several studies that predict noncompliance with standards governing working conditions (Short et al. 2016), food safety (Ibanez and Toffel 2020), oil wells (Mani and Muthulingam 2019), environmental performance and labor conditions (Villena et al. 2021), and air pollution (Schneider 2012, Bennett et al. 2013, Duflo et al. 2013, Pierce and Toffel 2013). Violations cited might differ from true violations because auditors might fail to detect and document some violations—the phenomenon we are specifically seeking to investigate. Like other studies of auditor bias using large real-world auditing datasets, we observe reported violations but not any violations that are missed. We address this challenge in several ways described in the Identification section below, including by examining systematic variation in a given establishment's violations when its auditor changes between second- and third-party auditors, and by exploiting the quasi-random assignment of auditors to audits—that is, random assignment conditional on controls we observe.

We have no reason to believe that *over*reporting violations is a problem. Auditors need to provide evidence to support their citation of violations, such as records documenting wage violations or photos documenting blocked aisles or lack of fire extinguishers. It would be time-consuming and difficult to fabricate this evidence in the course of a brief audit with factory managers and workers watching. In addition, factory managers must sign off on the auditor's report, stating that they recognize these

violations exist and will take actions to remediate them. Moreover, in our many conversations with representatives of global brands, auditors, factories, and labor advocates, we have heard many concerns about the integrity of auditors and their susceptibility to corruption, but none about violation fabrication.

*Under*reporting violations, however, is a substantial concern, in part because identifying and citing violations requires substantial effort and suppliers often seek to hide them. Auditors in Gujarat, India, for example, systematically underreported plant emissions as being just below the standard, "although true emissions were typically higher" (Duflo et al. 2013: 1499). As a result, following prior work on auditor and inspector bias (e.g., Duflo et al. 2013, Short et al. 2016, Ibanez and Toffel 2020), we interpret positive estimated coefficients on independent variables (i.e., effects that increase the number of violations cited, conditional on the controls described below) as a more accurate representation of the actual code-of-conduct violations and thus of higher audit quality.

3.3. Independent Variables

We created *second-party auditor* as a dummy variable coded 1 when an audit was conducted by second-party auditors (the global buyer's employees) and 0 when conducted by third-party auditors.

Testing H2, H3, and H4 requires measuring variables at the regional level. For the 10 countries in our sample that contain the most audits (96% of audits and 95% of factories)—China, India, Bangladesh, Turkey, the United Kingdom, Pakistan, Vietnam, Romania, Cambodia, and Myanmar—we use "region" to refer to the factory's state or province and, for the remaining 4% of audits (to maintain statistical power), to the factory's country. This resulted in 169 regions in our estimation sample. We measured a region's reported corruption (H2) as *regional reported bribery* at the region-year level using "bribery incidence (percent of firms experiencing at least one bribe payment request)" from World Bank Enterprise Survey data (http://www.enterprisesurveys.org). This dataset provides a rare source of panel data on subnational region-level corruption; our metric has been used by others to measure regional corruption (e.g., Cuervo-Cazurra 2016, Lee et al. 2020, Adbi et al. 2022).[7] To be clear, this does not measure corruption or bribery within the auditing regime we study, but instead assesses the prevalence of bribery reported in the region by respondents to the World Bank survey. To measure the pervasiveness of second-party auditing at the region-year level (H3), we created *regional prevalence of second-party auditing* as the proportion of a given region-year's audits conducted by second-party auditors.

Three binary variables classify audits according to whether there was audit firm rotation, lead

[7] For the United Kingdom, which is not included in this data, we used annual average values among Western European countries. When a region lacked bribery data at the beginning of our sample, we used the value from the first subsequent year for which bribery data were available for that region (which occurred for 66 regions in 16 countries, accounting for 1322 observations). When bribery data were not available for other region-year observations, we used the region's value from the prior year (which occurred for 111 regions in 25 countries, accounting for 7082 observations).

auditor rotation,[8] or neither. Firm rotation is indicated by coding *different audit firm* (H4) as 1 when the focal audit was conducted by a firm other than the one that conducted that factory's prior audit, and 0 otherwise. *Different lead auditor (but same audit firm)* is coded 1 when an audit was conducted by a different lead auditor but the same audit firm that conducted the factory's prior audit (i.e., individual lead auditor rotation), and 0 otherwise. *Same lead auditor (and same audit firm)* is coded 1 when an audit was conducted by same the lead auditor and audit firm as the factory's prior audit (i.e., no rotation), and 0 otherwise.[9] We use *different lead auditor (but same audit firm)* as the omitted (baseline) category in our regression models because we want *different audit firm* (H4) to distinguish the effects on audit quality of rotating audit firms from those of rotating lead auditors from the same firm.

3.4 Control variables

We include several controls in the specification to ensure that our empirical approach tests our hypothesized effects and rules out alternative explanations. Our model includes three sets of fixed effects. Supplier *factory* fixed effects control for the factory's industry, location, time-invariant elements of its institutional context, and other time-invariant factory-specific effects, such as general willingness or ability to comply with the code of conduct. *Audit-year* fixed effects control for secular changes that might affect factory conditions or auditor stringency. *Audit sequence*—whether the audit is the factory's first audit, second, and so on—accounts for the possibility that being audited provides factories the opportunity to learn what are considered violations and how to comply. We create a dummy variable for each audit-sequence value after first winsorizing (top-coding) the sequence count at its 95th percentile (eighth audit) to mitigate the influence of outliers, following an approach used by others (e.g., Bird et al. 2019).

We include additional audit-level variables to control for factors that might affect the number of violations cited. *Unannounced audit* equals 1 when the audited factory had no advanced notice of the audit and 0 otherwise. Receiving advanced notice provides time to resolve or conceal potential violations, which might reduce the number of violations cited.

Audit scope, which could affect the number of violations cited, is designated by four binary variables: (a) *new factory audit* refers to an initial audit conducted at a factory before the global buyer places an order; (b) *limited scope audit* refers to an audit focused on a new addition to a factory or on a new supplier factory whose prior audit was incomplete;[10] (c) *follow-up audit* refers to an audit typically

[8] We focus on lead auditors because our data names them but not any other members of their teams.

[9] We include in all models a dummy flagging the 70 audits in which the audit firm differed from the factory's prior audit but the lead auditor's name remained the same. This could result from different lead auditors with the same name or from a lead auditor switching employers. If we exclude these 70 observations as a robustness test, the results are nearly identical.

[10] For example, auditors may need to return to a new supplier factory if it does not have all documentation prepared

conducted a few months after an audit has yielded a "Red" or "Red-Critical" rating; and (d) *re-audit* refers to a routine audit conducted roughly annually after the factory's new-factory audit or another re-audit. These are mutually exclusive categories; for a given audit, only one of these variables is coded 1 and the others are 0. We include in our model specification *limited scope audit* and *follow-up audit* for both the factory's focal (current) audit and its prior (lagged) audit and *new factory audit* for the prior audit (new factory audits are not included as focal audits, given our use of lagged variables). We use *re-audit*—for both focal and prior audits—as the omitted categories.[11]

Because auditor gender can influence the number of violations cited (Short et al. 2016) and our data lacks lead auditor gender but includes their first name, we include a continuous variable, *female lead auditor*, that reflects the probability that a lead auditor's first name is a female name in the audited factory's country.[12] (We have no such data on other audit team members.) We used genderize.io, which uses profile information across major social networks to predict gender based on first names and has been used in other scholarship for this purpose (e.g., Cui et al. 2022, Luo and Zhang 2022).

We measured two factory-level characteristics that might influence the number of violations cited. *Percent factory supplied to global buyer* is the proportion of a factory's output supplied to FashionCo at the time of the focal audit, which might influence how beholden it was to FashionCo and how motivated to comply with FashionCo's Supplier Code of Conduct (Wilhelm and Villena 2021). *Workers (log)* is the natural log of the number of factory workers (plus 1) at the time of the audit, as factories with more employees might have more violations. (Because our specification includes factory fixed effects, which control for time-invariant values of factory output for this buyer and factory size, the coefficients on *percent factory supplied to global buyer* and *workers (log)* are properly interpreted as measuring how their variation over time affects factory violations.)[13]

We measured two annual institutional features that might influence factories' compliance with labor codes of conduct: the level of economic development and the extent of press freedom in the audited establishment's country, both measured at the country-year level. We controlled for a supplier factory country's level of economic development using annual values of the log of *GDP per capita* in US dollars, obtained from the World Bank.[14] *Press freedom index*, obtained from Reporters without Borders,

---

on their first visit.

[11] Because our model uses lagged independent variables that result in our estimation sample excluding each factory's first audit, no focal audits are *new factory audits*.

[12] We coded *female lead auditor* 0 for those audits for which the lead auditor's name could not be classified by gender and included in our models a separate dummy to indicate these observations.

[13] For each of the variables with missing values, we recoded these missing values to 0 (for binary variables) or as the variable's mean (for continuous variables) and included in our models a separate dummy indicating observations for which such recoding occurred. (See Table 3 notes.) Re-estimating our models but excluding these audits (as a robustness test) yielded broadly similar results.

[14] World Bank, "GDP per capita (current US$)," https://data.worldbank.org/indicator/NY.GDP.PCAP.CD, accessed

incorporates the threats journalists face, censorship and self-censorship, and the prevalence of journalists imprisoned, detained, murdered, attacked, or threatened.[15] We reverse-coded and rescaled this index so that higher scores on a 0-to-1 scale represent more press freedom (Toffel et al. 2015).

Tables 1 and 2 report summary statistics and correlations, respectively.

4. Empirical Analysis

Our setting provides a quasi-experiment with "as if random" assignment (Dunning 2008) of auditors to audits, which can be estimated using panel data regression methods. Simply put, after controlling for factors that could affect assignment or the number of violations (e.g., audit scope), our empirical approach compares audits of the same type and of the same factory over time, controlling for factors that could influence the actual number—not just the reported number—of violations. Thus, our analysis looks for statistical patterns to identify factors that lead to unusually high or low violation counts that correspond to our independent variables.

4.1. Empirical model

We estimate factory noncompliance recorded in an audit as a function of the hypothesized independent variables, controlling for the aforementioned factors that might affect the factory's on-the-ground working conditions. This approach mirrors that of prior studies of auditor and inspector reporting accuracy (e.g., Pierce and Snyder 2008, Schneider 2012, Bennett et al. 2013, Duflo et al. 2013, Pierce and Toffel 2013, Short et al. 2016, Ibanez and Toffel 2020). Specifically, we estimate a conditional fixed-effects Poisson model in which the conditional expectation $E(\cdot \mid \cdot)$ of noncompliance is:

$$E(Y_{id}|X_{id}, \alpha_i, \delta_y, \gamma_s, \lambda_{id}, \rho_{i,d-1}, \epsilon_{id}) = exp(\beta_1 X_{id} + \alpha_i + \delta_y + \gamma_s + \lambda_{id} + \rho_{i,d-1} + \epsilon_{id}). \quad (1)$$

Y_{id} refers to the *violations cited* in focal audit *d* of factory *i*. X_{id} represents our hypothesized explanatory variables: *second-party auditor* (H1), *regional reported bribery* (H2), *regional prevalence of second-party auditing* (H3), and *different audit firm* (H4).

The specification also includes the fixed effects and controls described above, including fixed effects for supplier factory *i* (α_i), audit year *y* (δ_y), and audit sequence value *s* (γ_s). λ_{id} represents the control variables associated with focal audit *d* of factory *i*: *unannounced audit*, *limited scope audit*, *follow-up audit, female lead auditor, percent factory supplied to global buyer,* and *workers (log)* as well as the *GDP per capita (log)* and *press freedom index* of the factory's country that year. It also includes *same lead auditor (and same audit firm)* to control for one form of audit rotation status (beyond *different audit firm*), and thus the reference (omitted) category for audit rotation is *different lead auditor (but same*

March 2023.

[15] Reporters Without Borders, "Detailed methodology: 2013–2021 editions," https://rsf.org/en/index-methodologie-2013-21?year=2018&data_type=general, accessed March 2023.

audit firm) that reflects individual—but not firm—rotation. $\rho_{i,d-1}$ refers to factors associated with the prior audit of factory *i*: *new factory audit (prior audit), limited scope audit (prior audit)*, and *follow-up audit (prior audit).* The focal (and prior) audit versions of *re-audit* are the omitted category for the focal (and prior) audit versions of the audit scope variables (*new factory audit, limited scope audit*, and *follow-up audit*). As described below, we use different estimation subsamples to test our hypotheses.

4.2. Identification

The ideal approach to test our hypotheses comparing second- to third-party auditors would be an experiment that randomizes the assignment of audits to second- and third-party auditors, making the assignment orthogonal to audited establishments' characteristics, including their industry, size, location, and working conditions. The audit assignment process in our context, though not purely random, is quasi-random; that is, random after conditioning on some observed factors. Specifically, our interviews with FashionCo employees revealed that, starting from the list of audits to be conducted, FashionCo typically prioritized assigning second-party auditors to new supplier factories before placing orders with them (these initial audits are excluded from our estimation sample, as described above) and to audits conducted a few months after an audit has yielded a "Red" or "Red-Critical" rating (our specifications include an indicator to flag such *follow-up audits*). The remaining audits were assigned amongst second- and third-party auditors to minimize an auditor's travel time between factories and to rotate each factory across auditors so that the second-party auditors would periodically visit each establishment. Rotating a factory's audits between second- and third-party auditors was meant to ensure that problematic working conditions would not go unreported by third-party auditors for too long. Our interviews indicated that, beyond the observable parameters described above, no other factors were considered. This assignment process, described to us by FashionCo's management and schedulers, was confirmed through our interviews at FashionCo's field offices and with factory personnel. (In the following subsections, we supplement this evidence from interviews with empirical analysis.)

Exploiting this quasi-random allocation of auditors to audits and the panel structure of our data—factories being audited by different auditors over time—our identification strategy relies on using (a) factory fixed effects to control for time-invariant characteristics of factories (including time-invariant country characteristics and each factory's average size and compliance level) and (b) a host of time-varying controls, and then assessing how changes between second- and third-party auditors affect the number of violations cited at the same factory. We conduct several preliminary analyses to assess our identification strategy and find evidence consistent with the quasi-random assignment process explained above.

4.2.1. Second-party versus third-party assignment (H1). To assess whether our data comports with our interviewees' assertions that assignment was random after conditioning on the observed factors

we control for, we estimate a regression model that predicts whether FashionCo was more likely to assign *second-party* auditors to factories whose prior audits yielded many violations (*violations cited (prior audit)*) and thus might be expected to do so in the focal audit; we henceforth refer to these as "worse factories." We include in this model all the right-hand–side variables from our main specification to test whether the allocation of auditors to audits is consistent with being random *conditional* on the observables included in our analyses. We cluster standard errors by country because auditor assignment decisions are made at the country level. As shown in Model 1 of Table A-1 in the Electronic Companion, the regression results yield a near-zero nonsignificant coefficient on *violations cited (prior audit)* (β = 0.0009, S.E. = 0.0015, p = 0.54), consistent with the quasi-random allocation our interviewees described. Thus, we find no evidence that second-party auditors were more (or less) likely to be assigned to audits of worse factories and therefore no evidence of this identification threat to H1.

4.2.2. Regional reported bribery (H2). H2 predicts that increases in regional reported corruption will be associated with a greater decrease in audit quality for third-party auditors than for second-party auditors. Empirical support for H2 would manifest as a widening gap in the violations cited by second-party versus third-party audits as regional reported corruption increases. The key identification threat to that analysis is that this pattern might also be observed if, as reported corruption increased within a region, third-party auditors were increasingly likely to be assigned to "better factories"—those at which prior audits yielded few violations and thus the focal audit might be expected to do so. Equivalently, this scenario would imply that second-party auditors would be more likely to be assigned to "worse factories" as regional reported corruption increases.

To assess this identification threat, we empirically examine whether second-party auditors were indeed more likely to be assigned to worse factories as regional reported bribery increased. We estimate a model that predicts whether an audit is assigned to a *second-party* auditor based on the interaction between the factory's *violations cited (prior audit)* and *regional reported bribery*. We control for both variables' direct effects as well as the other right-hand–side variables from our main specification. The results, reported as Model 2 in Table A-1 in the Electronic Companion, yield a near-zero nonsignificant coefficient on *violations cited (prior audit) X regional reported bribery* (β = -0.0001, S.E. = 0.0001, p = 0.24), which provides no evidence that FashionCo tended to increasingly assign third-party auditors to better factories as regional reported bribery increased, which alleviates concern about this identification threat to H2.

4.2.3. Regional presence of second-party auditors (H3). H3 predicts that third-party-auditor quality will be higher in regions with more second-party auditors. Empirical support would manifest as third-party audits citing more violations as the prevalence of second-party auditors in the region increases. The key identification threat is that this pattern might also be observed if, as the prevalence of second-

party auditors increased in a region, third-party auditors were more likely to be assigned to worse factories. (Equivalently, this would imply that second-party auditors would be more likely to be assigned to better factories as the prevalence of second-party auditors increases.) Our FashionCo interviewees indicated that this was not the case; rather, they established or increased the number of second-party auditors in regions in which the company's purchasing volume made it worthwhile to bear the overhead of doing so.

We nonetheless empirically examine whether FashionCo was indeed more likely to assign third-party auditors to worse factories as the prevalence of second-party auditors in a region increased. To do so, we estimate a regression model akin to the one used to assess the H2 identification threat. Here, we predict the probability that an audit is assigned to a *second-party* auditor based on the interaction between violations cited in the factory's prior audit and the *regional prevalence of second-party auditing*. We control for both variables' main effects and also include all other right-hand–side variables from our main specification. The near-zero nonsignificant coefficient on the *violations cited (prior audit)* × *regional prevalence of second-party auditing* interaction term ($\beta = 0.0025$, S.E. = 0.0050, $p = 0.63$) provides no evidence that the regional prevalence of second-party auditing affected the likelihood of third-party auditors being assigned to worse factories (results reported in Model 3 of Table A-1 in the Electronic Companion) and assuage concern over this identification threat to H3.

4.2.4. Third-party audit firm rotation (H4). H4 predicts that firm rotation among third-party auditors increases audit quality, manifested by more violations being reported. A key identification threat is the possibility that worse factories triggered such rotation—that is, they would be especially likely to be assigned to a different third-party audit firm. If so, that assignment process—rather than our hypothesized quality mechanism—could manifest as auditor rotation being associated with more violations. To assess this, we regressed *different audit firm* on *violations cited (prior audit)*, controlling for all the right-hand–side variables from our main specification (except for the audit firm rotation variables *different audit firm* and *same lead auditor (and same audit firm)*, given that this model is predicting audit rotation). Here, as when we test H4 in our main analysis, we isolate rotation among third-party auditors by restricting our sample to the subset of observations in which a factory's focal audit and prior audit were both conducted by third-party auditors. The results, reported in Model 4 of Table A-1 in the Electronic Companion, yield a small negative and nonsignificant coefficient on *violations cited (prior audit)* ($\beta = -0.0033$, S.E. = 0.0027, $p = 0.23$), which provides no evidence of this identification threat to H4.

4.3. Results

We estimate the models testing our hypotheses using quasi-maximum likelihood estimation of the Poisson likelihood function, which yields consistent estimates even for dependent variables that are not

Poisson-distributed provided that the conditional mean is correctly specified (Cameron and Trivedi 2010). We test H1 by estimating a model on the entire sample, clustering standard errors by factory because second- versus third-party auditors are assigned to factories, and report results in Model 1 of Table 3. The statistically significant positive coefficient on *second-party auditor* (Model 1: $\beta = 0.091$, SE = 0.019, $p < 0.01$, IRR = 1.10) indicates that second-party auditors record, on average, 10% more violations than third-party auditors. This finding supports H1, which predicts that second-party auditors outperform third-party auditors. While third-party audits in this sample record an average of 8.75 violations, the average marginal effect (AME) of 0.85 indicates that second-party auditors record an average of 9.60, controlling for all other factors in the model.

Overall, the control variables yield coefficients consistent with prior research. Audits conducted by the same lead auditor who conducted the factory's prior audit report 1.21 fewer violations than audits led by a different lead auditor from the same audit firm (the omitted category), consistent with studies extolling the benefits of rotating lead auditors of the same audit firm (Lennox 2014, Lennox et al. 2014, Short et al. 2016, Ball et al. 2017). Audits with a *female lead auditor* report 0.39 more violations than those with male lead auditors, directionally consistent with prior research showing that audit teams with at least one woman report more violations than all-male audit teams (Short et al. 2016). *Unannounced audits* yield 0.88 more violations than pre-announced audits, which is directionally in line with prior studies (Lebaron and Lister 2015, Short et al. 2016). As expected due to their more circumscribed scope, *follow-up audits* and *limited scope audits* yield fewer violations than *re-audits*, the omitted category. We find no evidence that the number of violations is affected by (a) within-factory (given the factory fixed effects) variation in *workers (log)* covered by the audit, (b) within-factory variation in *percent factory supplied to global buyer*, or (c) within-country variation of *GDP per capita (log)* or *press freedom index.*[16]

To test H2, we estimate our model on two subsamples—all focal audits by second-party auditors (Model 2) and all focal audits by third-party auditors (Model 3)—and compare the coefficients of *regional reported bribery*. We cluster standard errors by region in both models because reported perceptions of corruption are measured at the regional level. Estimating the model on the sample of focal audits conducted by second-party auditors yields a near-zero nonsignificant coefficient on *regional reported bribery* (Model 2: $\beta = -0.002$, SE = 0.005, $p = 0.73$), providing no evidence that second-party–audit results are affected by changes in regional reported bribery. In contrast, estimating the model on the sample of focal audits conducted by third-party auditors yields a statistically significant negative coefficient on *regional reported bribery* (Model 3: $\beta = -0.018$, SE = 0.006, $p < 0.01$, IRR = 0.98), indicating that third-party auditors record fewer violations as regional reported bribery increases. The

[16] Our inclusion of factory fixed effects limits these variables so as to identify only within-factory variation during our sample's few years, which might not provide sufficient power to detect effects.

AME indicates that third-party audits report 0.16 fewer violation for each unit increase in *regional reported bribery*. A Wald test indicates that these coefficients statistically differ (χ^2=5.39, p=0.02), which supports H2.

To test H3, we restrict our attention to Model 3, which is estimated only on focal audits by third-party auditors, with standard errors clustered by region because the prevalence of second-party auditors is measured by region. The positive, statistically significant coefficient on *regional prevalence of second-party auditing* (Model 3: $\beta = 0.378$, SE = 0.159, $p < 0.05$, IRR = 1.46) supports H3, which predicts higher third-party–auditor quality in regions with more second-party auditing. The AME of 3.34 indicates that for each 10-percentage-point increase (0.10) in the proportion of a region's audits conducted by second-party auditors, third-party auditors in that region record 0.33 more violations on average.

We test H4 by estimating our model on the subsample consisting of audits in which third-party auditors conducted a factory's focal and prior audits. This subsample contains three auditor-rotation scenarios: focal audits conducted by (a) the same third-party firm and the same lead auditor that conducted the prior audit (no rotation); (b) the same third-party firm as the prior audit but different lead auditor (third-party lead auditor rotation), which is the omitted category, and (c) a different third-party firm than that of the prior audit (audit firm rotation). Model 4 reports the results, with standard errors clustered by factory because audit firm rotation varies across a factory's audits. H4 is supported by the statistically significant positive coefficient on the variable *different audit firm* (Model 4: $\beta = 0.125$, SE = 0.021, $p < 0.01$, IRR = 1.13). This indicates that audits associated with third-party audit firm rotation (Scenario 3) cite 13% more violations—or slightly more than one additional violation, given the AME of 1.02—than audits associated with the third-party-lead-auditor–rotation scenario (the Scenario 2 baseline). For context, the statistically significant negative coefficient on the variable *same lead auditor (and same audit firm)* (Model 4: $\beta = -0.106$, SE = 0.036 , $p < 0.01$, IRR = 0.90, AME = -0.87) indicates that no-rotation third-party audits cite 0.87 fewer violations than audits conducted by a different lead auditor from the same third-party firm that conducted the factory's prior audit; this confirms prior research findings that lead-auditor rotation prompts higher-quality audits. Combining these two effects, our results indicate that, compared to the no-rotation audits (Scenario 1), audits by a different third-party audit firm (Scenario 3) record nearly two additional violations (1.02+0.87=1.89, a 23% increase with respect to this sample's mean of 8.1 violations).

While the magnitude of some of our marginal effects is small, in this empirical context small effects can have big consequences. Detecting even a single violation that would otherwise remain unreported can be valuable. Billions of dollars per year are spent on social audits seeking to detect working condition violations of suppliers' codes of conduct. Audits report an average of 9.2 violations in our sample at a given factory. Detecting even just one additional violation per audit would appreciably

increase the number of violations reported in terms of both percentage at the individual factory and in the magnitude of the aggregate across all audits conducted every day around the world. In addition, reporting any additional violations has a snowball effect because it provides more opportunities to prompt corrective actions such as training or installing machine guards that can both prevent recurrences of the specific violations detected and prevent other types of violations from occurring.

4.4. Robustness Tests and Extensions

Our results are robust to several alternative analyses. First, while our primary approach to addressing the 70 audits that had the same lead auditor name but a different audit firm as the factory's prior audit was to flag them with an indicator variable, our results are nearly identical when we instead exclude those audits from our sample.

Second, our primary approach to addressing missing values for the audit or bribery variables was to recode them (to 0 for dummies and to means for continuous variables) and to include in our models a series of dummies flagging such recoding. Our results are broadly similar when we instead re-estimate our models excluding audits with missing values.

Third, we re-estimated our primary models using OLS regression instead of Poisson regression, after changing the dependent variable from the number of *violations cited* to the *log of violations cited* (after adding 1 to avoid missing values). These OLS models, which avoid dropping panels with all-zero violations and thus report slightly larger sample sizes, yield results, reported in Table A-2 in the Electronic Companion, consistent with our primary results reported in Table 3.

Fourth, we explored an alternative approach to testing H2 by estimating Model 1 on two subsamples: "less-corrupt regions" and "more-corrupt regions" whose average reported corruption levels during our sample period were either below or else at or above the sample median. The results, reported in Table A-3 in the Electronic Companion, yield a near-zero non-significant coefficient on *second-party auditor* in the less-corrupt–region subsample (β=-0.010, SE=0.019, p=0.60) and a positive and marginally significant coefficient on *second-party auditor* in the more-corrupt–region subsample (β=0.112, SE=0.058, p=0.06). Critically, a Wald test indicates that the difference between these coefficients is statistically significant (Wald χ^2=3.93, $p<0.05$), suggesting—in support of H2—that second-party auditors are especially likely to outperform third-party auditors in regions with more reported bribery.

Finally, violations vary in criticality, which raises the question of whether auditors might respond differently to the incentives we hypothesize when faced with major versus minor violations, perhaps viewing major violations as a bigger threat that would be particularly risky to leave unreported. We note that global buyers care about all types of violations because even minor ones can have major consequences. An interview with a FashionCo employee revealed that while "one-off issues" were typically coded as minor violations, they could nonetheless have significant implications. For example,

finding one aisle in a factory blocked would be coded as a minor violation, even though it could trap workers in a fire. She added that factories that failed to address such minor violations risked escalating situations that would eventually be coded as major violations. Concern about minor violations and the desire to resolve them before they grow into major ones was shared by the auditors we shadowed; we heard several times from auditors that they like to "nip violations in the bud."

Nonetheless, the possibility that second- and third-party auditors might treat major and minor violations differently prompted us to conduct a supplementary analysis by reestimating the specification in Table 3 of Model 1 to predict *minor violations cited* and *major violations cited* separately instead of all *violations cited*. The results, reported in Table A-4 in the Electronic Companion, show that second-party auditors cite more of both types of violation, as evidenced by statistically significant positive coefficients on *second-party auditor* in both our models that predict *minor violations cited* (Model 1: $\beta = 0.199$, SE = 0.059, $p<0.01$, IRR = 1.22) and *major violations cited* (Model 2: $\beta = 0.165$, SE = 0.036, $p < 0.01$, IRR = 1.18).

5. Discussion

In a world of complex, multi-layered business relationships and intricate business, regulatory, and social demands, firms must vigilantly monitor suppliers. Prevailing wisdom considers third-party auditors the gold standard, but our study complicates this assumption. We show that second-party auditors outperform third-party auditors in our context—on average and especially in more corrupt regions. We also demonstrate the efficacy of several strategies for improving third-party–audit quality, including sourcing from both internal and external providers and rotation across third-party audit firms.

Our sample is limited to a single brand and its supply chain and we do not claim that second-party auditors always outperform third-party auditors. Second-party auditors are especially likely to detect more violations when the global buyer employing them is committed to learning the truth about supplier misconduct. If, instead, the buyer wants to cover up such misconduct, that, too, could be accomplished more effectively using auditors who are employees. However, our finding that a leading global buyer exercises control of its auditors to generate more (rather than less) information about supplier misconduct is, in itself, an important contribution. It opens new conversations in a field in which stakeholders mistrust the motives of global buyers and thus reflexively demand third-party audits (e.g., Lebaron and Lister 2015). It also provides several important scholarly and managerial insights for designing and implementing better monitoring strategies, which we discuss below.

5.1. Academic Contributions

This research builds on and extends four research streams: quality management and monitoring, responsible supply chains and social auditing, bias and auditor quality, and behavioral operations.

5.1.1. Quality management and monitoring. We contribute to this literature by focusing on determinants of the quality of audits themselves, rather than on how auditing practices affect the quality of audited processes, products, or services. The operations literature characterizes poor supplier working conditions as a "process quality problem concerning how products are sourced and produced" (Chen et al. 2020: 1234) and research has theorized that auditing can affect process quality in this domain by incentivizing supplier evasion (Plambeck and Taylor 2016) and supplier-auditor collusion (Chen et al. 2020) as well as by depressing social responsibility performance in certain tiers of the supply chain under certain circumstances (Huang et al. 2022). Although the universe of quantitative research is limited due to the difficulty of obtaining audit data, empirical studies have documented several ways that auditing can improve process and service quality: companies learn from inspections (Mani and Muthulingam 2019); audits foment spillover of quality knowledge across shared suppliers (Muthulingam and Agrawal 2016); and external audits can improve retailers' on-shelf availability, inventory accuracy, and—ultimately—sales (Chuang et al. 2015). We extend this body of research by empirically examining how the quality of the *auditing process itself* can be improved. Although our data do not permit us to explore the relationship between audit quality and the quality of audited processes, our findings suggest new ways to improve process quality overall on the assumption that better quality auditing leads to better quality processes.

5.1.2. Responsible supply chains and social auditing. Our work also contributes to the literature on supply chain social responsibility and sustainability (e.g., Kraft et al. 2018, Lo et al. 2018, Buell and Kalkanci 2021) by focusing on a key but often overlooked operational tool: social auditing. Managers are increasingly expected to protect the environment (Muthulingam et al. 2022) and improve labor practices not only in their own organizations (Kesavan et al. 2022) but throughout their supply chains (Caro et al. 2021, Huang et al. 2022). Research has identified several mechanisms that can promote more responsible supplier practices: incentives such as financial benefits for socially responsible businesses (Kraft et al. 2018, Buell and Kalkanci 2021, Kesavan et al. 2022); institutional pressures from regulatory agencies, buying firms, and NGOs (Villena et al. 2021); and relational arrangements with global buyers, such as lean training and relational capital (Distelhorst et al. 2017, Villena et al. 2021). This interdisciplinary literature has also documented various determinants of suppliers' compliance with labor standards, including institutional pressures (Locke et al. 2007; Ang et al. 2012, Anner 2012, Anner et al. 2013, Distelhorst et al. 2015, Toffel et al. 2015, Bartley and Egels-Zandén 2016, Short et al. 2020), organizational structures (Distelhorst et al. 2017, Bird et al. 2019, Amengual et al. 2020), buyer purchasing practices (Oka 2010, Locke 2013), and social compliance program design features (Lindholm et al. 2016, Short et al. 2020, Distelhorst and Shin 2023). But it has paid little attention to the monitoring mechanisms that give global brand managers visibility into suppliers' social performance.

Social auditing is the operational lynchpin of many companies' social responsibility programs. It

is the basis for supplier accountability to global brands' contractual demands and normative expectations, and for global brands' accountability to stakeholders such as NGOs, activists, and shareholders. It will be the basis for compliance with the EU's Corporate Sustainability Due Diligence Directive, which starting in 2024 establishes a corporate due diligence duty on EU companies and many companies doing business in the EU to identify, prevent, and mitigate negative human rights and environmental impacts in their value chains. While other mechanisms could be used to promote corporate social responsibility—for instance, direct government regulation or private liability—the prevailing human rights accountability regime has been built around corporate due diligence, which depends to a great extent on social auditing. Within this regime, our findings and the tools we provide to improve social auditing quality can support companies' broader strategies to improve supply chain social responsibility and sustainability and to manage emerging regulatory risks.

Our expansion of the toolkit for addressing supply chain labor standards is especially valuable in light of recent research findings that institutional and incentive mechanisms shown to improve suppliers' *environmental* practices are *not* as effective for improving their *labor* practices (Villena et al. 2021). The tools we provide to help brands identify supplier violations of labor standards enable brands to better target interventions to address these issues. In addition, our findings suggest ways to build the kinds of relationships that have been found to improve supplier social responsibility. Research has shown that global brands can cascade sustainability practices throughout the supply chain by developing collaborative relationships with tier-one suppliers (Villena and Gioia 2018). Brands' use of their own second-party auditors for monitoring might help build such relationships by demonstrating a commitment to the supplier.

Our work thus contributes to the limited empirical research on supply chain auditors, which has focused exclusively on third-party auditor performance (Short et al. 2016; Short et al. 2020). Although there is deep skepticism of third-party auditors among scholars in this field (e.g., O'Rourke 2002, Egels-Zandén and Lindholm 2015, Lebaron and Lister 2015), to our knowledge, no study has empirically compared their performance to that of other types of auditors. Our study should provoke conversation about what combination of auditors—under what conditions—can most effectively foster improved working conditions in global supply chains.

5.1.3. Bias and auditor quality. This literature has focused on how bias and conflicts of interest affect audit quality. For instance, the accuracy of third-party audits is influenced by (a) who pays for them (Ronen 2010, Duflo et al. 2013, Short et al. 2016), (b) the level of competition among inspection firms (Bennett et al. 2013), and (c) incentives for audit firms to cross-sell to clients (Koh et al. 2013, Pierce and Toffel 2013). The policy prescription most clearly supported by this literature is simple: more auditor independence (McAllister 2012, Lebaron and Lister 2015). Our results complicate this account by

suggesting that independence is not a single, silver-bullet solution to improve auditing. While greater independence (via third-party audits) might help combat very real conflicts of interest that can bias auditors, it might simultaneously create agency problems that undermine audit quality—for example, by corner-cutting. We identify mechanisms for improving auditor performance that attend to both sets of incentives.

5.1.4. Behavioral operations. Our work contributes to the behavioral operations literature by identifying counterintuitive behavioral drivers of auditor performance that challenge conventional assumptions about auditors' incentives and reveal novel ways to improve their performance. Our finding that the incentive structure created by auditor independence can degrade performance complements research demonstrating that incentives often work in unexpected ways (Lee et al. 2022). For instance, Kesavan et al. (2014) show that, although temporary and part-time workers do—as conventionally theorized—give retailers flexibility that can increase productivity, overuse of such workers undermines productivity by changing workplace dynamics and introducing new training and integration costs. Other studies have shown that incentive structures can backfire when workers do not follow assigned schedules (Ibanez et al. 2018) or rely on workarounds (Tucker et al. 2020). Our results extend this body of research by demonstrating the complexity of the incentive structure facing social auditors and the behavioral distortions it can yield. Specifically, while the incentive structure created by auditor independence can mitigate conflicts of interest, it can also create agency problems that exacerbate corner-cutting. Recognizing such tension can lead to more nuanced and effective interventions to improve audit quality. In this way, we contribute to research by proposing novel interventions—such as scheduling (e.g., Ibanez and Toffel 2020)—to improve monitor performance. Auditor allocation and audit firm rotation can be viewed as forms of scheduling used not only to allocate resources across tasks, but also to adjust incentive systems to improve performance.

5.2. Managerial Implications

Our results have important implications for managers of global brands monitoring supplier performance and for managers of the many other firms at high risk of reputational spillovers from the poor performance of business partners such as franchisees, distributors, vendors, and purchasing agents.

First, our study challenges the prevailing wisdom that third-party audits are always the most reliable. Our findings confirm some managers' worries about the trend of over-relying on third-party audits and some managers' intuition that their own auditors can do better. FashionCo managers told us that they believe their second-party auditors are more effective because they "work only for us," they "have more equity in what they do," and they pursue "a more holistic approach to social compliance"—for example, by trying to identify root causes and potential solutions. Our findings support their perceptions. While outside stakeholders may perceive third-party audits as more legitimate than second-

party audits, managers should be aware that they might learn more from their own auditors about what is happening on the ground, which can ultimately provide better protection for workers, the environment, and the company's reputation.

Accordingly, managers should periodically evaluate their monitoring programs and work closely with third parties to close any performance gaps detected. Managers who choose second-party audits to gain information benefits might need to engage with stakeholders such as activists and NGOs—who often expect or demand third-party audits—about the benefits of different auditing arrangements. It may be in the interest of both business and society for companies to maintain second-party auditors in the regions where they operate to foster the competition benefits we identify. Our findings should also inform activists and NGO leadership in their engagement with brands to protect workers.

Second, our findings help managers mitigate the tradeoffs they face when second-party audits yield higher-quality information than third-party audits but are more expensive. We demonstrate that managers need not sacrifice the typical cost-saving benefits of outsourced audits to bolster auditor performance. Allocating second- and third-party audits strategically can significantly improve the quality of external audits. These findings provide a lever for managers who wish to capture the reliability benefits of second-party audits but lack the resources to insource all audits. For example, they could allocate more audits to third parties in regions with less reported corruption. Managers should also rotate not only lead auditors but also auditing firms (including their own) to ensure that their auditors do not settle into routines that may cause them to overlook important information.

5.3. Limitations and Future Research

We acknowledge several limitations to our study. First, as with many other studies reliant on global supply chain factory audit data (e.g., Locke et al. 2007, Amengual et al. 2020), our data come from a single firm. We use the best data available (obtained via a data-sharing agreement) to test our hypotheses. Our data covers (a) factories manufacturing products for FashionCo and many other retailers and (b) auditors from many of the major auditing firms used by leading global buyers. FashionCo deliberately sought out third-party companies that audited for other major brands, so our sample is representative of auditing companies used by many global buyers. Future research could assess how the effects we identify extend to second-party auditors of different global buyers and could establish more circumstances that improve third-party audit quality beyond the three we identify (lower regional reported corruption, more second-party auditors in the region, and audit firm rotation).

Second, the proprietary audit dataset we obtained does not allow us to control for auditor training or professional credentials, which have been shown to affect audit results (Short et al. 2016). While we cannot rule out the possibility that training and skill affected audit results in ways we cannot observe, we note that (a) the second- and third-party auditors in our sample received comparable training on the

buyer's code of conduct, (b) the buyer works only with reputable multinational third-party auditors, and (c) our direct observations of audits did not expose differences in audit skill.

Third, as in most empirical studies of auditing, data limitations allow us to theorize—but not empirically test—the underlying mechanisms driving auditor performance. For instance, we cannot discern the extent to which performance gaps might be due to third-party auditors exerting less effort or receiving less transparent information from suppliers—perhaps due to fears that the auditors might disclose information to other clients. It is also possible that suppliers are instead less transparent with second-party auditors, who are closer to the managers making sourcing decisions.

Fourth, as with other auditing and inspection research (e.g., Pierce and Snyder 2008, Schneider 2012, Bennett et al. 2013, Duflo et al. 2013, Pierce and Toffel 2013, Short et al. 2016, Ibanez and Toffel 2020), we as researchers cannot observe actual conditions. Consistent with prior studies, we rely on regression specifications that include many controls to look for the presence of remaining variation in reported violations that correlates with our hypothesized variables; the absence of a logical connection between our hypothesized variables and the number of *actual* violations leads us to infer that any significant correlations between our hypothesized variables and the number of *reported* violations reflect auditor bias. Moreover—also consistent with prior work (e.g., Duflo et al. 2013, Short et al. 2016, Ibanez and Toffel 2020)—we concern ourselves with auditors underreporting violations rather than overreporting them (i.e., fabricating reports of violations) because citing violations requires recording supporting evidence. Our findings thus create opportunities for further study, including investigating what combinations of second- and third-party audits promote effective monitoring of other types of business partners besides suppliers, such as franchisees, distributors, vendors, and purchasing agents. Future research could also drill down into the mechanisms we theorize; for instance, testing auditor accuracy under varying reputation-risk conditions, such as reputation-damaging events or mandatory disclosure of audit results. Studies could also investigate the efficacy of different strategies to close quality gaps between auditors. Finally, future research should explore the extent to which our findings apply in other monitoring contexts, such as environmental compliance and food safety.

6. Conclusion

Seeking to benefit from the superior credibility and flexibility and lower cost of external assessments, many global buyers are relying less on their own (second-party) auditors and more on third-party auditors to monitor their supply chains. Some observers worry that this trend risks eroding audit quality. Data from one of the world's largest fashion brands confirms that these concerns are warranted. We find that the brand's own second-party auditors cite more violations than third-party auditors do and that this quality gap is wider in regions where reported corruption is higher, suggesting where global buyers

should be especially wary of relying on third-party auditors. We reveal two strategies that appear to improve the accuracy of third-party auditors: increasing the proportion of a region's audits conducted by second-party auditors and rotating the firms that audit a particular supplier. These auditor-allocation strategies can help global buyers manage brand risk and drive more sustainable and socially responsible practices in their supply chains.

References

Adbi A, Liu X, Mishra A (2022) Technology licensing and productivity growth: Evidence from manufacturing firms in developing economies. *Manufacturing & Service Operations Management* 2(1): 214–234.

Al-Thuneibat AA, Al Issa RTI, Baker RAA (2011) Do audit tenure and firm size contribute to audit quality? *Managerial Auditing Journal* 26(4): 317–334.

Amengual M, Distelhorst G, Tobin D (2020) Global purchasing as labor regulation: The missing middle. *ILR Review* 73(4): 817–840.

Ammenberg J, Wik G, Hjelm O (2001) Auditing external environmental auditors—Investigating how ISO 14001 is interpreted and applied in reality. *Eco-Management and Auditing: The Journal of Corporate Environmental Management* 8(4): 183–192.

Anderson E, Schmittlein DC (1984) Integration of the sales force: An empirical examination. *RAND Journal of Economics* 15(3): 385–395.

Ang D, Brown D, Dehejia R, Robertson R (2012) Public disclosure, reputation sensitivity, and labor law compliance: Evidence from Better Factories Cambodia. *Review of Development Economics* 16(4): 594–607.

Anner M (2012) Corporate social responsibility and freedom of association rights: The precarious quest for legitimacy and control in global supply chains. *Politics & Society* 40(4): 604–639.

Anner M, Bair J, Blasi J (2013) Towards joint liability in global supply chains: Addressing the root causes of labor violations in international subcontracting networks. *Comparative Labor Law and Policy Journal* 35(1): 1–43.

Ball G, Siemsen E, Shah R (2017) Do plant inspections predict future quality? The role of investigator experience. *Manufacturing & Service Operations Management* 19(4): 534–550.

Bartley T, Child C (2014) Shaming the corporation: The social production of targets and the anti-sweatshop movement. *American Sociological Review* 79(4): 653–679.

Bartley T, Egels-Zandén N (2016) Beyond decoupling: Unions and the leveraging of corporate social responsibility in Indonesia. *Socio-Economic Review* 14(2): 231–255.

Bennett VM, Pierce L, Snyder JA, Toffel MW (2013) Customer-driven misconduct: How competition corrupts business practices. *Management Science* 59(8): 1725–1742.

Bird Y, Short JL, Toffel MW (2019) Coupling labor codes of conduct and supplier labor practices: The role of internal structural conditions. *Organization Science* 30(4): 847–867.

Bradach JL (1997) Using the plural form in the management of restaurant chains. *Administrative Science Quarterly* 42(2): 276–303.

Brass DJ, Butterfield KD, Skaggs BC (1998) Relationships and unethical behavior: A social network perspective. *Academy of Management Review* 23(1): 14–31.

Buell RW, Kalkanci B (2021) How transparency into internal and external responsibility initiatives influences consumer choice. *Management Science* (67)2: 932–950.

Cameron, AC, Trivedi, PK (2010) *Microeconometrics Using Stata,* revised edition (Stata Press, College Station, TX).

Carcello, JV, Nagy, AL (2004) Audit firm tenure and fraudulent financial reporting. *Auditing: A Journal of Practice & Theory* 23(2): 55–69.

Caro F, Lane L, Sáez de Tejada Cuenca A (2021) Can brands claim ignorance? Unauthorized subcontracting in apparel supply chains. *Management Science* 67(4): 1993–2656.
Chen L, Yao S, Zhu K (2020) Responsible sourcing under supplier-auditor collusion. *Manufacturing & Service Operations Management* 22(6): 1234–1250.
China Labor Watch (2009) Corrupt audits damage worker rights: A case analysis of corruption in Bureau Veritas factory audits. Retrieved from https://core.ac.uk/download/pdf/5127032.pdf, accessed December 2022.
Chuang H, Oliva R, Liu S (2015) On-shelf availability, retail performance, and external audits: A field experiment. *Production and Operations Management* 25(5): 935–951.
Coffee JC Jr (2004) Gatekeeper failure and reform: The challenge of fashioning relevant reforms. *Boston University Law Review* 84(2): 301–364.
Cressey DR, Moore CA (1983) Managerial values and corporate codes of ethics. *California Management Review* 25(4): 53–77.
Cuervo-Cazurra A (2016) Corruption in international business. *Journal of World Business* 51(1): 35–49.
Cui R, Ding H, Zhu F (2022) Gender inequality in research productivity during COVID-19. *Manufacturing & Service Operations Management* 24(2): 707–726.
Distelhorst G, Hainmueller J, Locke RM (2017) Does lean improve labor standards? Management and social performance in the Nike supply chain. *Management Science* 63(3): 707–728.
Distelhorst G, Locke RM, Pal T, Samel H (2015) Production goes global, compliance stays local: Private regulation in the global electronics industry. *Regulation & Governance* 9(3): 224–242.
Distelhorst G, Shin JE (2023) Assessing the social impact of corporations: Evidence from management control interventions in the supply chain to increase worker wages. *Journal of Accounting Research* 61(3): 855–890.
Duflo E, Greenstone M, Pande R, Ryan N (2013) Truth-telling by third-party auditors and the response of polluting firms: Experimental evidence from India. *Quarterly Journal of Economics* 128(4): 1499–1545.
Dunning T. (2008) Improving causal inference: Strengths and limitations of natural experiments. *Political Research Quarterly* 61(2): 282–293.
Egels-Zandén N, Lindholm H (2015) Do codes of conduct improve worker rights in supply chains? A study of Fair Wear Foundation. *Journal of Cleaner Production* 107: 31–40.
Francis JR, Wilson ER (1988) Auditor changes: A joint test of theories relating to agency costs and auditor differentiation. *Accounting Review* 64(4): 663–682.
Heide JB, Kumar A, Wathne KH (2014) Concurrent sourcing, governance mechanisms, and performance outcomes in industrial value chains. *Strategic Management Journal* 35(8): 1164–1185.
Huang L, Song J-S, Swinney R (2022) Managing social responsibility in multitier supply chains. *Manufacturing & Service Operations Management* 24(6): 2797–3306.
Huq FA, Chowdhury IN, Klassen RD (2016) Social management capabilities of multinational buying firms and their emerging market suppliers: An exploratory study of the clothing industry. *Journal of Operations Management* 46(1): 19–37.
Ibanez MR, Clark JR, Huckman RS, Staats BR (2018) Discretionary task ordering: Queue management in radiological services. *Management Science* 64(9): 4389–4407.
Ibanez MR, Toffel MW (2020) How scheduling can bias quality assessment: Evidence from food-safety inspections. *Management Science* 66(6): 2396–2416.
Inditex (2017) Inditex Annual Report 2017. https://static.inditex.com/annual_report_2017/assets/pdf/memoria_en.pdf, accessed August 2022.
Inditex (2021) Inditex Annual Report 2021. https://static.inditex.com/annual_report_2021/en/documents/annual_report_2021.pdf, accessed August 2022.
Jensen MC, Meckling WH (1976) Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics* 3(4): 305–360.
Jia N (2018) The "make and/or buy" decisions of corporate political lobbying: Integrating the economic

efficiency and legitimacy perspectives. *Academy of Management Review* 43(2): 307–326.
Jin GZ, Lee J (2018) A tale of repetition: Lessons from Florida restaurant inspections. *Journal of Law and Economics* 61(1): 159–188.
John G, Weitz BA (1988) Forward integration into distribution: An empirical test of transaction cost analysis. *Journal of Law, Economics, and Organization* 4(2): 337–355.
Kesavan S, Lambert SJ, Williams JC, Pendem PK (2022) Doing well by doing good: Improving retail store performance with responsible scheduling practices at The Gap, Inc. *Management Science* 68(11): 7793–8514.
Kesavan S, Staats BR, Gilland W (2014) Volume flexibility in services: The costs and benefits of flexible labor resources. *Management Science* 60(8): 1884–1906.
King B, McDonnell M (2015) Good firms, good targets: The relationship among corporate social responsibility, reputation, and activist targeting. Tsutsui K, Lim A, eds. *Corporate Social Responsibility in a Globalizing World* (Cambridge University Press, New York), 430–454.
Klein B, Crawford RG, Alchian AA (1978) Vertical integration, appropriable rents, and the competitive contracting process. *Journal of Law and Economics* 21(2): 297–326.
Klotz D, Chatterjee K (1995) Dual sourcing in repeated procurement competitions. *Management Science* 41(8): 1317–1327.
Koh K, Rajgopal S, Srinivasan S (2013) Non-audit services and financial reporting quality: Evidence from 1978 to 1980. *Review of Accounting Studies* 18(1): 1–33.
Kouakou D, Boiral O, Gendron Y (2013) ISO auditing and the construction of trust in auditor independence. *Accounting, Auditing & Accountability Journal* 26(8): 1279–1305.
Kraakman RH (1986) Gatekeepers: The anatomy of a third-party enforcement strategy. *Journal of Law, Economics, & Organization* 2(1): 53–104.
Kraft T, Valdes L, Zheng Y (2018) Supply chain visibility and social responsibility: Investigating consumers' behaviors and motives. *Manufacturing & Service Operations Management* 20(4): 617–636.
Lebaron G, Lister J (2015) Benchmarking global supply chains: The power of the "ethical audit" regime. *Review of International Studies* 41(5): 905–924.
Lee C-C, Wang C-W, Ho S-J (2020) Country governance, corruption, and the likelihood of firms' innovation. *Economic Modelling* 92: 326–338.
Lee HS, Kesavan S, Kuhnen C (2022) When do group incentives for retail store managers work? *Production and Operations Management* 31(8): 3077–3095.
Lennox, CS (2014) Auditor tenure and rotation. Hay D, Knechel WR, Willekens M, eds. *The Routledge Companion to Auditing* (Routledge, New York), 89–106.
Lennox CS, Wu X, Zhang T (2014) Does mandatory rotation of audit partners improve audit quality? *Accounting Review* 89(5): 1775–1803.
Levitt A (2000) Renewing the covenant with investors. Speech delivered to Center for Law and Business, New York University, May 10, 2000. US Securities and Exchange Commission. https://www.sec.gov/news/speech/spch370.htm, accessed August 2020.
Lindholm H, Egels-Zandén N, Rudén C (2016) Do code of conduct audits improve chemical safety in garment factories: Lessons on corporate social responsibility in the supply chain from Fair Wear Foundation. *International Journal of Occupational and Environmental Health* 22(4): 283–291.
Ljungqvist A, Marston F, Starks LT, Wei KD, Yan H (2007) Conflicts of interest in sell-side research and the moderating role of institutional investors. *Journal of Financial Economics* 85(2): 420–456.
Lo CKY, Tang CS, Zhou Y, Yeung ACL, Fan D (2018) Environmental incidents and the market value of firms: An empirical investigation in the Chinese context. *Manufacturing & Service Operations Management* (20)3: 422–439.
Locke RM (2013) *The Promise and Limits of Private Power: Promoting Labor Standards in a Global Economy* (Cambridge University Press, Cambridge, UK).
Locke RM, Qin F, Brause A (2007) Does monitoring improve labor standards? Lessons from Nike. *ILR Review* 146(1–2): 21–40.

Luo H, Zhang L (2022) Scandal, social movement, and change: #MeToo in Hollywood. *Management Science* 68(2): 1278–1296.
Mani V, Muthulingam S (2019) Does learning from inspections affect environmental performance? Evidence from unconventional well development in Pennsylvania. *Manufacturing & Service Operations Management* (21)1: 177–197.
Manik JA, Yardley J (2012) Bangladesh finds gross negligence in factory fire. *New York Times,* Dec. 17. https://www.nytimes.com/2012/12/18/world/asia/bangladesh-factory-fire-caused-by-gross-negligence.html, accessed May 2021.
Mayer KJ (2006) Spillovers and governance: An analysis of knowledge and reputational spillovers in information technology. *Academy of Management Journal* 49(1): 69–84.
McAllister LK (2012) Regulation by third-party verification. *Boston College Law Review* 53(1): 1–32.
Merchant B (2017) Life and death in Apple's forbidden city. *The Guardian (UK)*, June 18. https://www.theguardian.com/technology/2017/jun/18/foxconn-life-death-forbidden-city-longhua-suicide-apple-iphone-brian-merchant-one-device-extract, accessed May 2021.
Metzger M, Dalton DR, Hill JW (1993) The organization of ethics and the ethics of organizations: The case for expanded organizational ethics audits. *Business Ethics Quarterly* 3(1): 27–43.
Muthulingam S, Agrawal A (2016) Does quality knowledge spillover at shared suppliers? An empirical investigation. *Manufacturing & Service Operations Management* 18(4): 525–544.
Muthulingam S, Dhanorkar S, Corbett CJ (2022) Does water scarcity affect environmental performance? Evidence from manufacturing facilities in Texas. *Management Science* 68(4): 2785–2805.
Narayanan S, Balasubramanian S, Swaminathan JM (2009) A matter of balance: Specialization, task variety, and individual learning in a software maintenance environment. *Management Science* 55(11): 1861–1876.
Nickerson JA, Silverman BS (2003) Why aren't all truck drivers owner-operators? Asset ownership and the employment relation in interstate for-hire trucking. *Journal of Economics & Management Strategy* 12(1): 91–118.
Nike (2018) Maximum performance minimum impact: FY16/17 NIKE, Inc. Sustainable Business Report. https://about.nike.com/en/newsroom/resources/reports, accessed August 2022.
Nike (2022) Barriers: FY21 NIKE, Inc. Impact Report. https://about.nike.com/en/newsroom/resources/reports, accessed August 2022.
Norton SW (1988) An empirical look at franchising as an organizational form. *Journal of Business* 61(2): 197–218.
Oka C (2010) Channels of buyer influence and labor standard compliance: The case of Cambodia's garment sector. *Advanced Industrial Labor Relations* 17: 153–183.
Oliva R, Sterman JD (2001) Cutting corners and working overtime: Quality erosion in the service industry. *Management Science* 47(7): 894–914.
O'Rourke D (2002) Monitoring the monitors: A critique of PricewaterhouseCooper's labor monitoring. Jenkins R, Pearson R, Seyfang G, eds. *Corporate Responsibility and Ethical Trade: Codes of Conduct in the Global Economy* (Earthscan: London), 196–207.
Parmigiani A (2007) Why do firms both make and buy? An investigation of concurrent sourcing. *Strategic Management Journal* 28(3): 285–311.
Pierce L, Snyder J (2008) Ethical spillovers in firms: Evidence from vehicle emissions testing. *Management Science* 54(11): 1891–1903.
Pierce L, Toffel MW (2013) The role of organizational scope and governance in strengthening private monitoring. *Organization Science* 24(5): 1558–1584.
Plambeck EL, Taylor TA (2016) Supplier evasion of a buyer's audit: Implications for motivating supplier social and environmental responsibility. *Manufacturing & Service Operations Management* 18(2):184–197.
Prakash A, Potoski M (2007) Collective action through voluntary environmental programs: A club theory perspective. *Policy Studies Journal* 35(4): 773–792.
Puranam P, Gulati R, Bhattacharya S (2013) How much to make and how much to buy? An analysis of

optimal plural sourcing strategies. *Strategic Management Journal* 34(10): 1145–1161.
Ronen J (2010) Corporate audits and how to fix them. *Journal of Economic Perspectives* 24(2): 189–210.
Rubin PH (1978) The theory of the firm and the structure of the franchise contract. *Journal of Law & Economics* 21(1): 223–233.
Rushe D (2016) Retail group approves Bangladesh factories as safety concerns persist, report finds. *The Guardian (UK),* Nov. 21. https://www.theguardian.com/world/2016/nov/21/bangladesh-garment-factories-safety-alliance-rana-plaza-report, accessed May 2021.
Schneider HS (2012) Agency problems and reputation in expert services: Evidence from auto repair. *Journal of Industrial Economics* 60(3): 406–433.
Short JL, Toffel MW, Hugill AR (2016) Monitoring global supply chains. *Strategic Management Journal* 37(9): 1878–1897.
Short JL, Toffel MW, Hugill AR (2020) Improving working conditions in global supply chains: The role of institutional environments and monitoring program design. *ILR Review* 73(4): 873–912.
Spicer A (2021) Why New Zealand needs a Modern Slavery Act. *Consumer NZ* (May 27) https://www.consumer.org.nz/articles/modern-slavery, accessed July 2022.
Tepalagul N, Lin L (2015) Auditor independence and audit quality: A literature review. *Journal of Accounting, Auditing & Finance* 30(1): 101–121.
Toffel MW, Short JL, Ouellet M (2015) Codes in context: How states, markets, and civil society shape adherence to global labor standards. *Regulation & Governance* 9(3): 205–223.
Trevino LK, Butterfield KD, McCabe DL (1998) The ethical context in organizations: Influences on employee attitudes and behaviors. *Business Ethics Quarterly* 8(3): 447–476.
Tucker AL, Zheng S, Gardner JW, Bohn RE (2020) When do workarounds help or hurt patient outcomes? The moderating role of operational failures. *Journal of Operations Management* 66(1–2): 67–90.
US Securities and Exchange Commission (2003) Strengthening the commission's requirements regarding auditor independence. 17 CFR 210, 240, 249, and 274 (March 27). https://www.sec.gov/rules/final/33-8183.htm, accessed June 2020.
Vanpoucke E., Klassen RD (Forthcoming) Reducing forced labour in supply chains: what could traditional companies learn from social enterprises? *International Journal of Operations & Production Management*.
Villena VH, Gioia DA (2018) On the riskiness of lower-tier suppliers: Managing sustainability in supply networks. *Journal of Operations Management* 64(1): 65–87.
Villena VH, Wilhelm M, Xiao C-Y (2021) Untangling drivers for supplier environmental and social responsibility: An investigation in Philips Lighting's Chinese supply chain. *Journal of Operations Management* 67(4): 476–510.
Wilhelm M, Villena VH (2021) Cascading sustainability in multi-tier supply chains: When do Chinese suppliers adopt sustainable procurement? *Production and Operations Management* 30(11): 4198–4218.
Yang Z, Aydin G, Babich V, Beil DR (2012) Using a dual-sourcing option in the presence of asymmetric information about supplier reliability: Competition vs. diversification. *Manufacturing & Service Operations Management* 14(2): 202–217.

Table 1: Summary Statistics

	Mean	SD	Min	Max
Violations cited	9.19	7.05	0	74
Second-party auditor	0.27	0.44	0	1
Regional reported bribery (region-year)	19.45	16.72	0	82.10
Regional prevalence of second-party auditing (region-year)	0.36	0.21	0	1
Rotation				
Different audit firm	0.57	0.50	0	1
Same lead auditor (and same audit firm)	0.15	0.36	0	1
Different lead auditor (but same audit firm) (omitted category)	0.28	0.45	0	1
Audit sequence (top-coded at 95th percentile)	4.22	1.96	2	8
Unannounced audit	0.23	0.42	0	1
Female lead auditor	0.38	0.45	0	1
Audit scope				
Re-audit (omitted category)	0.52	0.50	0	1
Limited scope audit	0.10	0.30	0	1
Follow-up audit	0.32	0.47	0	1
Missing (thus unknown)	0.05	0.22	0	1
Prior audit's scope				
Re-audit (prior audit) (omitted category)	0.41	0.49	0	1
New factory audit (prior audit)	0.14	0.35	0	1
Limited scope audit (prior audit)	0.09	0.28	0	1
Follow-up audit (prior audit)	0.24	0.43	0	1
Missing (thus unknown)	0.12	0.33	0	1
Percent factory supplied to global buyer	0.33	0.25	0.00	1
Workers (log)	5.42	1.23	1.39	9.27
GDP per capita (log)	8.44	0.96	6.32	11.16
Press freedom index	0.61	0.17	0.00	0.97

Notes: The unit of observation is a factory audit. N=10,078 observations.

Table 2. Correlations

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
(1) Violations cited	1.00															
(2) Second-party auditor	0.10	1.00														
(3) Regional reported bribery	0.21	0.17	1.00													
(4) Regional prevalence of second-party auditing	0.17	0.43	0.36	1.00												
(5) Different audit firm	0.11	-0.14	-0.01	0.07	1.00											
(6) Same lead auditor (and same audit firm)	-0.06	0.18	0.13	0.05	-0.48	1.00										
(7) Unannounced audit	0.13	0.18	0.17	0.24	0.04	0.01	1.00									
(8) Female lead auditor	-0.04	-0.08	-0.21	-0.19	-0.03	0.00	-0.12	1.00								
(9) Limited scope audit	-0.01	0.33	0.04	0.03	-0.16	0.15	-0.06	0.03	1.00							
(10) Follow-up audit	0.04	0.04	0.11	0.06	0.04	-0.06	0.13	-0.07	-0.23	1.00						
(11) New factory audit (prior audit)	-0.03	0.11	-0.03	-0.04	-0.09	0.09	-0.07	0.04	0.46	-0.23	1.00					
(12) Limited scope audit (prior audit)	0.05	0.03	0.02	0.00	0.03	-0.01	0.02	-0.01	0.21	-0.18	-0.12	1.00				
(13) Follow-up audit (prior audit)	0.05	-0.02	0.10	0.04	0.07	-0.08	0.08	-0.07	-0.16	0.16	-0.23	-0.17	1.00			
(14) Percent factory supplied to global buyer	-0.02	0.04	0.09	0.06	0.02	0.03	0.05	-0.06	-0.02	0.04	-0.02	-0.05	0.06	1.00		
(15) Workers (log)	0.19	0.12	0.55	0.23	-0.04	0.10	0.16	-0.14	0.00	0.08	-0.05	0.00	0.08	-0.12	1.00	
(16) GDP per capita (log)	-0.17	-0.23	-0.81	-0.48	-0.03	-0.09	-0.16	0.31	-0.03	-0.16	0.06	-0.01	-0.14	-0.18	-0.50	1.00
(17) Press freedom index	0.11	0.04	-0.11	0.14	0.10	-0.14	-0.03	0.02	0.04	-0.01	0.05	0.07	0.06	-0.05	-0.08	0.10

Notes: The unit of observation is a factory audit. N=10,078 observations.

Table 3: Poisson Regression Results (Dependent Variable: *Violations cited*)

		(1)		(2)		(3)		(4)	
	Sample:	Entire sample		Focal audits by second-party auditors		Focal audits by third-party auditors		Focal and prior audits by third-party auditors	
	Standard error clustering:	Factory		Region		Region		Factory	
		Coef [SE]	AME	Coef [SE]	AME	Coef [SE]	AME	Coef [SE]	AME
H1	**Second-party auditor**	**0.091*** [0.019]**	**0.85**						
H2	**Regional reported bribery**	-0.015*** [0.004]	-0.14	**-0.002 [0.005]**	**-0.02**	**-0.018*** [0.006]**	**-0.16**	-0.016*** [0.004]	-0.13
H3	**Regional prevalence of second-party auditing**	0.098 [0.077]	0.91	-0.031 [0.178]	-0.33	**0.378** [0.159]**	**3.34**	0.296*** [0.115]	2.43
H4	**Different audit firm**	0.105*** [0.015]	0.97	-0.005 [0.027]	-0.05	0.117*** [0.022]	1.04	**0.125*** [0.021]**	**1.02**
	Same lead auditor (and same audit firm)	-0.130*** [0.026]	-1.21	-0.196*** [0.031]	-2.05	-0.082** [0.036]	-0.73	-0.106*** [0.036]	-0.87
	Unannounced audit	0.095*** [0.018]	0.88	0.072*** [0.026]	0.75	0.100*** [0.029]	0.88	0.078** [0.035]	0.64
	Female lead auditor	0.042*** [0.015]	0.39	0.105*** [0.036]	1.09	0.044 [0.034]	0.39	0.084*** [0.022]	0.69
	Limited scope audit	-0.273*** [0.038]	-2.53	-0.316*** [0.051]	-3.29	-0.202*** [0.049]	-1.79	-0.163** [0.072]	-1.34
	Follow-up audit	-0.136*** [0.016]	-1.26	-0.238*** [0.039]	-2.48	-0.126*** [0.019]	-1.11	-0.130*** [0.024]	-1.07
	New factory audit (prior audit)	-0.038 [0.042]	-0.35	-0.351*** [0.108]	-3.65	0.036 [0.049]	0.32	0.048 [0.070]	0.39
	Limited scope audit (prior audit)	-0.094*** [0.029]	-0.87	-0.203*** [0.076]	-2.11	-0.069* [0.041]	-0.61	-0.041 [0.057]	-0.33
	Follow-up audit (prior audit)	-0.072*** [0.016]	-0.66	-0.073** [0.037]	-0.76	-0.090*** [0.020]	-0.79	-0.083*** [0.023]	-0.68
	Percent factory supplied to global buyer	0.004 [0.048]	0.04	0.120 [0.117]	1.25	0.008 [0.066]	0.07	-0.011 [0.075]	-0.09
	Workers (log)	0.024 [0.021]	0.22	0.078* [0.046]	0.81	0.010 [0.030]	0.08	0.007 [0.029]	0.06
	GDP per capita (log)	-0.139 [0.123]	-1.29	0.082 [0.187]	0.86	0.059 [0.219]	0.52	0.108 [0.202]	0.89
	Press freedom index	0.198* [0.104]	1.83	0.078 [0.271]	0.81	0.131 [0.235]	1.16	-0.280* [0.154]	-2.29
	Factory fixed effects	Yes		Yes		Yes		Yes	
	Audit-year fixed effects	Yes		Yes		Yes		Yes	
	Audit-sequence fixed effects	Yes		Yes		Yes		Yes	
	Observations (audits)	10,078		2,001		7,038		5,093	
	Factories	2,313		729		1,848		1,421	
	Regions	169		52		157		147	
	Countries	34		13		34		33	
	Sample mean of dependent variable	9.2		10.7		8.6		8.1	

Notes: Poisson coefficients with robust standard errors (in brackets) clustered by factory for Models 1 and 4 and by region for Models 2 and 3. * p<0.10, ** p<0.05, *** p<0.01. AME is average marginal effect. The unit of observation is a factory audit. All models include a series of dummy variables (not shown) indicating instances in which missing values from audits were recoded to (a) 0 for *female lead auditor* (12% of the 10,087 observations in our main sample), *unannounced audit* (2%), audit scope (*limited scope, follow-up, re-audit*; 5%), and prior audit scope (*new factory, limited scope, follow-up, re-audit*; 12%) or (b) the variable's mean for *percent factory supplied to global buyer* (26%) and *workers* (4%). All models include a dummy indicating audits conducted by a lead auditor whose name is the same as that in the prior audit but for which the auditing firm rotated, which could be a coincidence (two people with the same name) or the lead auditor changing firms (1%). *Different lead auditor (but same audit firm)* is the omitted category for *different audit firm* (H4) and *same lead auditor (and same audit firm). Re-audit* is the omitted category for the audit scope variables (*new factory audit, limited scope audit*, and *follow-up audit*).

Electronic Companion to
Second- versus Third-party Audit Quality:
Evidence from Global Supply Chain Monitoring

Table A-1: Identification Threat Tests

	(1)	(2)	(3)	(4)
Identification threat test for:	H1: Are second-party auditors assigned to "worse factories"?	H2: Are third-party auditors increasingly assigned to "better factories" as regional corruption increases?	H3: Are third-party auditors increasingly assigned to "better factories" as regional prevalence of second-party auditing increases?	H4: Is third-party audit firm rotation especially likely for "worse factories"?
Dependent variable:	Second-party auditor	Second-party auditor	Second-party auditor	Different audit firm
Sample:	Entire sample	Entire sample	Entire sample	Focal and prior audits by third-party auditors
Standard error clustering:	Country	Country	Country	Country
Violations (prior audit)	**0.0009**	0.0034*	-0.0001	**-0.0033**
	(0.0015)	(0.0019)	(0.0021)	**(0.0027)**
Regional reported bribery	-0.0031	-0.0020	-0.0032*	0.0045**
	(0.0019)	(0.0023)	(0.0018)	(0.0022)
Violations (prior audit) X Regional reported bribery		**-0.0001**		
		(0.0001)		
Regional prevalence of second-party auditing	0.7567***	0.7609***	0.7313***	0.2499*
	(0.0409)	(0.0401)	(0.0656)	(0.1286)
Violations (prior audit) X Regional prevalence of second-party auditing			**0.0025**	
			(0.0050)	
Different audit firm	-0.0598***	-0.0598***	-0.0597***	
	(0.0209)	(0.0210)	(0.0210)	
Same lead auditor (and same audit firm)	0.0643***	0.0652***	0.0642***	
	(0.0214)	(0.0217)	(0.0213)	
Unannounced audit	0.1480***	0.1477***	0.1482***	-0.0470
	(0.0515)	(0.0512)	(0.0514)	(0.0368)
Female lead auditor	-0.0229	-0.0233	-0.0226	-0.0218
	(0.0204)	(0.0208)	(0.0206)	(0.0160)
Limited scope audit	0.4363***	0.4358***	0.4362***	-0.0873
	(0.0464)	(0.0462)	(0.0465)	(0.0863)
Follow-up audit	0.0871***	0.0855***	0.0875***	-0.0249**
	(0.0177)	(0.0179)	(0.0179)	(0.0099)
New factory audit (prior audit)	0.0354	0.0361	0.0357	-0.0957
	(0.0364)	(0.0362)	(0.0363)	(0.0699)
Limited scope audit (prior audit)	-0.0119	-0.0113	-0.0122	0.0680*
	(0.0257)	(0.0253)	(0.0251)	(0.0370)
Follow-up audit (prior audit)	0.0165	0.0160	0.0166	0.0198*
	(0.0144)	(0.0141)	(0.0144)	(0.0117)
Percent factory supplied to global buyer	0.0406	0.0421	0.0405	-0.0347
	(0.0318)	(0.0323)	(0.0320)	(0.0344)
Workers (log)	0.0294***	0.0295***	0.0292***	-0.0220
	(0.0092)	(0.0091)	(0.0092)	(0.0203)
GDP per capita (log)	-0.1880***	-0.1912***	-0.1884***	0.6451***
	(0.0483)	(0.0494)	(0.0487)	(0.2032)
Press freedom index	0.1488***	0.1459***	0.1496***	-0.2140**
	(0.0506)	(0.0496)	(0.0496)	(0.0970)
Factory fixed effects	Yes	Yes	Yes	Yes
Audit-year fixed effects	Yes	Yes	Yes	Yes
Audit-sequence fixed effects	Yes	Yes	Yes	Yes
Observations (audits)	10,078	10,078	10,078	5,508
Factories	2313	2313	2313	1836
Regions	169	169	169	161
Countries	34	34	34	34

Notes: OLS regression coefficients with robust standard errors clustered by country in parentheses; *** p<0.01, ** p<0.05, * p<0.10. The unit of observation is a factory audit. Model 4 excludes the rotation variables (*different audit firm* and *same lead auditor (and same audit firm)*) because the dependent variable in that model is *different audit firm*. All models include a series of dummy variables (not shown) indicating instances in which missing values from audits were recoded to (a) 0 for *female lead auditor*, *unannounced audit*, audit scope (*limited scope, follow-up, re-audit*), and prior audit scope (*new factory, limited scope, follow-up, re-audit*) or (b) the variable's mean for *percent factory supplied to global buyer* and *workers*. All models include a dummy indicating audits conducted by a lead auditor whose name is the same as that in the prior audit but for which the auditing firm rotated, which could be a coincidence (two people with the same name) or the lead auditor changing firms. *Different lead auditor (but same audit firm)* is the omitted category for *different audit firm* and *same lead auditor (and same audit firm). Re-audit* is the omitted category for the audit scope variables (*new factory audit, limited scope audit*, and *follow-up audit*).

Table A-2: Re-estimating Primary Models Using OLS Regression to Predict Log of *Violations Cited* (after adding 1)

Dependent variable: Violations cited, logged after adding 1

		(1)	(2)	(3)	(4)
	Sample:	Entire sample	Entire sample	Focal audits by second-party auditors	Focal audits by third-party auditors
	Standard error clustering:	Factory	Region	Region	Factory
H1	**Second-party auditor**	**0.074*** **			
		(0.018)			
H2	**Regional reported bribery**	-0.009**	**-0.002**	**-0.015****	-0.017***
		(0.004)	**(0.005)**	**(0.006)**	(0.004)
H3	**Regional prevalence of second-party auditing**	0.197***	-0.027	**0.427*****	0.350***
		(0.071)	(0.142)	**(0.135)**	(0.106)
H4	**Different audit firm**	0.107***	0.003	0.117***	**0.123*****
		(0.015)	(0.029)	(0.017)	**(0.020)**
	Same lead auditor (and same audit firm)	-0.114***	-0.166***	-0.088***	-0.094***
		(0.022)	(0.025)	(0.028)	(0.033)
	Unannounced audit	0.095***	0.073**	0.094***	0.109***
		(0.017)	(0.028)	(0.035)	(0.033)
	Female lead auditor	0.020	0.081*	0.027	0.047**
		(0.015)	(0.041)	(0.032)	(0.022)
	Limited scope audit	-0.239***	-0.317***	-0.165***	-0.134**
		(0.035)	(0.049)	(0.049)	(0.065)
	Follow-up audit	-0.132***	-0.189***	-0.132***	-0.130***
		(0.015)	(0.038)	(0.014)	(0.021)
	New factory audit (prior audit)	-0.037	-0.245**	-0.000	-0.008
		(0.039)	(0.101)	(0.041)	(0.063)
	Limited scope audit (prior audit)	-0.081***	-0.182**	-0.050	0.051
		(0.028)	(0.071)	(0.042)	(0.053)
	Follow-up audit (prior audit)	-0.064***	-0.048	-0.081***	-0.077***
		(0.015)	(0.038)	(0.018)	(0.021)
	Percent factory supplied to global buyer	0.005	0.118	-0.004	0.017
		(0.046)	(0.108)	(0.061)	(0.070)
	Workers (log)	0.042**	0.060*	0.044**	0.035
		(0.019)	(0.032)	(0.020)	(0.029)
	GDP per capita (log)	0.107	0.156	0.220	0.275
		(0.120)	(0.171)	(0.186)	(0.188)
	Press freedom index	0.222**	0.232	0.144	-0.016
		(0.092)	(0.202)	(0.192)	(0.127)
	Factory fixed effects	Yes	Yes	Yes	Yes
	Audit-year fixed effects	Yes	Yes	Yes	Yes
	Audit-sequence fixed effects	Yes	Yes	Yes	Yes
	Observations (audits)	11,033	3,075	7,958	5,798
	Factories	3,256	1,803	2,756	2,114
	Regions	200	104	193	180
	Countries	38	18	38	38
	Sample mean of dependent variable	9.3	10.4	8.8	8.2

Notes: OLS coefficients with robust standard errors (in parentheses) clustered by factory for Models 1 and 4 and by region for Models 2 and 3. * p<0.10, ** p<0.05, *** p<0.01. See Table 3 for additional specification notes.

Table A-3: Robustness Test of H2, Estimating Model 1 of Table 3 on Low versus High Regional Reported Corruption

Dependent variable: Violations cited

	(1)		(2)	
Sample:	Low regional corruption (<sample median)		High regional corruption (>= sample median)	
Standard error clustering:	Region		Region	
	Coef [SE]	AME	Coef [SE]	AME
Second-party auditor	**-0.010**	**-0.08**	**0.112***	**1.07**
	(0.019)		**(0.058)**	
Regional prevalence of second-party auditing	0.408**	3.17	0.109	1.05
	(0.208)		(0.146)	
Different audit firm	0.113***	0.88	0.101***	0.97
	(0.035)		(0.023)	
Same lead auditor (and same audit firm)	-0.095**	-0.74	-0.132***	-1.27
	(0.042)		(0.033)	
Unannounced audit	0.058	0.45	0.101***	0.97
	(0.039)		(0.026)	
Female lead auditor	0.026	0.20	0.049	0.47
	(0.016)		(0.035)	
Limited scope audit	-0.274**	-2.13	-0.262***	-2.51
	(0.109)		(0.036)	
Follow-up audit	-0.057**	-0.44	-0.156***	-1.50
	(0.028)		(0.016)	
New factory audit (prior audit)	0.009	0.07	-0.045	-0.43
	(0.072)		(0.040)	
Limited scope audit (prior audit)	-0.058	-0.45	-0.087***	-0.84
	(0.073)		(0.030)	
Follow-up audit (prior audit)	-0.027	-0.21	-0.084***	-0.80
	(0.021)		(0.020)	
Percent factory supplied to global buyer	-0.078	-0.61	0.045	0.43
	(0.068)		(0.047)	
Workers (log)	0.130***	1.01	0.009	0.08
	(0.025)		(0.026)	
GDP per capita (log)	0.409	3.18	-0.183	-1.76
	(0.472)		(0.201)	
Press freedom index	0.089	0.69	0.087	0.84
	(0.244)		(0.230)	
Factory fixed effects	Yes		Yes	
Audit-year fixed effects	Yes		Yes	
Audit-sequence fixed effects	Yes		Yes	
Observations (audits)	2,286		7,820	
Factories	577		1,763	
Regions	87		93	
Countries	17		28	
Sample mean of dependent variable	7.8		9.6	

Notes: Poisson coefficients with robust standard errors (in parentheses) clustered by region. * p<0.10, ** p<0.05, *** p<0.01. See Table 3 for additional specification notes. A Wald test indicates that the coefficients on *second-party auditor* across the two models statistically differ (Wald χ^2=3.93, p<0.05).

Table A-4: Predicting Minor and Major Violations

	(1)	(2)
Dependent variable:	Minor violations cited	Major violations cited
Sample:	Entire sample	Entire sample
Standard error clustering:	Factory	Factory
Second-party auditor	**0.199*** **	**0.165*** **
	(0.059)	**(0.036)**
Regional reported bribery	-0.008	-0.013***
	(0.008)	(0.005)
Regional prevalence of second-party auditing	-0.133	0.077
	(0.176)	(0.117)
Different audit firm	0.092*	0.274***
	(0.051)	(0.028)
Same lead auditor (and same audit firm)	-0.491***	-0.372***
	(0.069)	(0.049)
Unannounced audit	-0.123*	0.153***
	(0.064)	(0.032)
Female lead auditor	0.256***	0.137***
	(0.057)	(0.029)
Limited scope audit	-0.407***	-0.729***
	(0.122)	(0.063)
Follow-up audit	-0.355***	-0.489***
	(0.055)	(0.030)
New factory audit (prior audit)	-0.223*	-0.135*
	(0.124)	(0.070)
Limited scope audit (prior audit)	0.116	-0.098**
	(0.096)	(0.048)
Follow-up audit (prior audit)	0.035	-0.092***
	(0.059)	(0.029)
Percent factory supplied to global buyer	0.192	-0.006
	(0.146)	(0.079)
Workers (log)	0.128*	0.067*
	(0.069)	(0.039)
GDP per capita (log)	-0.652**	0.252
	(0.316)	(0.175)
Press freedom index	-1.259***	0.138
	(0.258)	(0.127)
Factory fixed effects	Yes	Yes
Audit-year fixed effects	Yes	Yes
Audit-sequence fixed effects	Yes	Yes
Observations (audits)	7,905	10,000
Factories	1,604	2,283
Regions	155	163
Countries	31	31
Sample mean of dependent variable	0.8	4.0

Notes: These models decompose the prediction of all *violations cited*—from Model 1 of Table 3—to separately predict *minor violations cited* and *major violations cited*. Poisson coefficients with robust standard errors (in brackets) clustered by factory. * p<0.10, ** p<0.05, *** p<0.01. See the paper's Table 3 for additional notes.