HARVARD | BUSINESS | SCHOOL

Auditor Independence and Outsourcing: Aligning Incentives to Mitigate Shilling and Shirking

Ashley Palmarozzo
Jodi L. Short
Michael W. Toffel

Working Paper 21-078

Auditor Independence and Outsourcing: Aligning Incentives to Mitigate Shilling and Shirking

Ashley Palmarozzo
Harvard Business School

Jodi L. Short
University of California

Michael W. Toffel
Harvard Business School

Working Paper 21-078

Funding for this research was provided in part by Harvard Business School. We thank Primark, and especially Graham Clewer, for providing access to its supplier code of conduct audit data.

Auditor Independence and Outsourcing: Aligning Incentives to Mitigate Shilling and Shirking[†]

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

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

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

January 6, 2021

Multinational corporations (MNCs) hire auditors to assess their business partners' compliance with quality, working conditions, and environmental standards. Independent third-party auditors are widely assumed to outperform second-party auditors employed and thus controlled by MNCs. Synthesizing literatures on auditor independence and outsourcing decisions, we compare how independence and control can affect auditor performance. Using proprietary data from a global apparel brand, we find that second-party auditors outperform independent third-party auditors, and that third-party auditors' performance improves when MNCs concurrent source audits, using both second- and third-party auditors. However, both second- and third-party auditors perform better with more independence from the entities they audit—specifically, when auditing factories most recently audited by a different firm. These findings yield important insights for more effective monitoring of business partners.

Keywords: concurrent sourcing; transaction cost economics; outsourcing; suppliers; monitoring

INTRODUCTION

In many business and regulatory contexts, firms use auditors to monitor and assess the quality of difficult-to-observe aspects of their business partners' performance. Companies use auditors to ensure that their franchisees, suppliers, distributors, purchasing agents, and other business partners do not shirk on contractual terms or performance quality. They rely on auditors to assess their partners' adherence to management system standards such as the ISO 9001 quality standard, the ISO 14001 environmental standard, and the ISO 26000 human rights standard. Firms often engage in monitoring to manage the risk of reputational spillovers from their

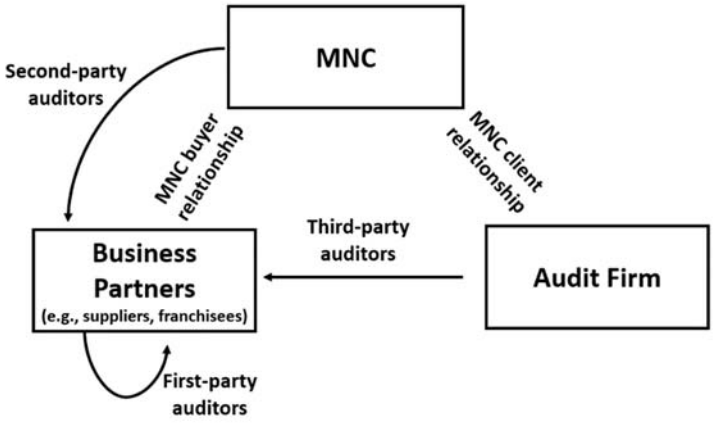
[†] We thank Primark, and especially Graham Clewer, for providing access to its supplier code of conduct audit data. We gratefully acknowledge the research assistance of Melissa Ouellet and helpful feedback from Xiang Ao, Hong Luo, Feng Zhu, and participants of the Alliance for Research on Corporate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

business partners' misconduct (Heide, Kumar, and Wathne, 2014). Some companies monitor business partners in regulated contexts ranging from anti-corruption to pharmaceutical distribution to protect themselves from vicarious liability for their business partners' illegal activities. In these and many similar assurance regimes, the task of verifying compliance with specified standards is typically outsourced to independent third-party auditors.

Auditor independence is a central feature of organizational and governmental assurance regimes. Indeed, according to some, independence "is the very essence that gives an auditor's work its value" (Levitt, 2000) and is what makes third-party auditing "the gold standard" (Prakash and Potoski, 2007: 790). Independence is conceptualized as *separation* between the auditor and the audited entity, operationalized as a relationship in which they have no employment, investment, or other formal financial ties. Independence is designed to mitigate a specific problem: shilling, or auditors covering up the wrongdoing of an audited firm so as to mislead others into doing business with the firm. In theory, the separation associated with independence enhances the auditor's performance—their ability to comprehensively and accurately assess the audited entity's conduct—by insulating the auditor from the bias that might afflict assessments conducted by employees of the audited firm (first-party auditors), who are beholden to that firm for their livelihood (Coffee, 2004). Thus, it is widely assumed that assessments by external, third-party auditors are more credible and legitimate than firms' own assessments of their business partners using their own internal auditors (second-party audits), because second-party auditors may believe that their employer gains a competitive advantage from partners' misconduct. Figure 1 illustrates the relationships among (a) companies seeking to monitor their business partners, (b) auditors, and (c) audited business partners—and the terms we use to describe these relationships.

However, this basic conception of independence providing separation that safeguards auditors from bias has serious limitations. First, several studies demonstrate that third-party auditors are not free from bias even when they meet formal independence criteria. For instance, third-party auditors are more incentivized to overlook poor performance when they are directly compensated by the audited entity (Duflo et al., 2013; Jiang, Stanford, and Xie, 2012; Pierce and Toffel, 2013), when they have opportunities to cross-sell to the audited entity (Causholli et al., 2014; Pierce and Toffel, 2013), or when they audit the same entity multiple times (Griffin and Tang, 2011; Lennox, Wu, and Zhang, 2014; Short, Toffel, and Hugill, 2016).

Figure 1: Business Partner Monitoring



Second, the model of independence as separation assumes a dyadic relationship between the auditor and the audited entity, ignoring the multi-layered nature of business relationships in which many assurance regimes are embedded. This is probably due to the fact that much auditing research focuses on financial auditing, in which there are indeed two key players: the outside auditor and the audited firm. But contemporary business relationships can be much more complex, particularly those that extend across global value chains in which multinational companies (MNCs) outsource production to a network of suppliers—often through sourcing

agents—and many first-tier suppliers subcontract some of the work to second- and third-tier suppliers. The concept of independence as separation provides little guidance about how to structure monitoring in such multi-layered relationships. For instance, the third-party auditors typically hired by MNCs to monitor their suppliers are formally independent of these supply chain partners. But so are the MNC's own employees, who could conduct second-party audits. Independence, on its own, does not tell MNCs whether business partner monitoring will be performed better by insourced second-party auditors or by outsourced third party auditors.

Third, while conceptualizing independence as separation targets some of the issues that can undermine auditor performance, it ignores others. Apart from the problem of shilling, auditors might shirk—cutting corners to save costs or increase output. Like shilling, shirking can cause auditors to overlook critical information about the audited firm and produce incomplete and thus misleading audit reports. But independence is not designed to address shirking. Indeed, separation between the auditor and the audited entity might actually incentivize it because MNCs may find it more difficult to monitor third-party auditors than second-party ones. This suggests the need to identify alternative mechanisms for raising auditor performance.

This paper tackles these puzzles by synthesizing insights from the literatures on auditor independence and outsourcing decisions, including literature on transaction cost economics (TCE) and concurrent sourcing, to theorize several factors that enhance or undermine auditor performance. We propose a novel conceptualization of auditor performance as a function not only of independence, but also of incentives created by the choice to either insource or outsource monitoring. We explore these questions in the context of auditing global supply chain factories for compliance with labor standards contained in MNCs' supplier codes of conduct. In response to pressure from activists and nongovernmental organizations (NGOs), thousands of MNCs,

including all Fortune 500 companies, have adopted these codes of conduct that require their suppliers to meet specified workplace standards (McBarnet, 2007), and many MNCs conduct audits to monitor and assess suppliers' adherence to those codes. This empirical context is particularly useful for exploring these issues because it entails complex and multi-layered relationships among buyers, suppliers, and auditors in which firms engage in extensive business partner monitoring to manage business, legal, and reputational risks.

We develop and test our hypotheses using a proprietary auditing database from a global branded apparel MNC that faces the choice of monitoring suppliers' compliance with its labor code of conduct by using its own employees as second-party auditors and/or by outsourcing to third-party auditors. We refer to the MNC henceforth as "the MNC buyer" when discussing its relationship with its suppliers and as "the MNC client" when referring to its relationship with third-party auditors. First, we develop hypotheses that offer competing predictions about mechanisms promoting high-quality auditor performance: independence or control. We find, after controlling for factors to account for actual violations present, that second-party auditors outperform third-party auditors by recording more new code-of-conduct violations, suggesting the limitations of client-auditor independence as the touchstone of auditor performance.

Second, we hypothesize that a concurrent sourcing strategy—in which MNCs source audits of factories in a given market from both second-party and third-party auditors—affects the incentives of third-party auditors that will improve their performance. We find that it does.

Third, beyond focusing on the independence of the MNC client-to-auditor relationship, we explore how independence between auditors and the suppliers they monitor corresponds to auditor performance. Specifically, we hypothesize that the rotation of audit companies will improve auditor performance, above and beyond the known benefits of rotating individual

auditors. We find that it does for both third-party and second-party auditors.

These results have important theoretical and managerial implications. In terms of theory, our study extends auditor independence research by treating auditor performance quality as an issue not only of independence but also of sourcing strategy. We extend the outsourcing literature by empirically examining whether sourcing strategies theorized in the literature to be superior actually lead to performance differences across insourced and outsourced production. We also provide robust evidence of the effect of audit firm rotation, an important question on which the auditing literature has reached little consensus.

Our study generates several managerial insights that can help firms strategically monitor business partners. First, whereas industry wisdom holds that third-party auditors are key to high-quality auditor performance (Lebaron and Lister, 2015), we show that, at least in our sample, MNCs' own employees—serving as second-party auditors—perform better. Second, we demonstrate that firms can improve third-party auditing performance by having second-party auditors conduct some of the audits in the same market. Third, MNCs can better monitor business partners by rotating monitoring firms.

LITERATURE REVIEW

Auditor independence is achieved by outsourcing monitoring to third parties, yet the auditor independence literature has not conceptualized auditor selection as an outsourcing decision. Thus, there has been little productive dialogue between the literature on auditor independence and literatures on outsourcing such as TCE and concurrent sourcing. Perhaps this is unsurprising, since these strands of literature have very different empirical focus and theoretical motivation. Auditing research tends to assume that the principal threat to audit quality is shilling—auditors covering up audited entities' misbehavior—and therefore focuses on

independence as the principal mechanism to control that threat. Consequently, these studies emphasize the relationship between performance outcomes and various factors that might influence the impartiality of third-party auditors, such as incentives, social and institutional conditions, and cognitive biases (e.g., Duflo et al., 2013; Kumar and Charkrabarti, 2012; Short, Toffel, and Hugill, 2016). By contrast, the TCE and concurrent sourcing literatures focus on the strategic decision of how to structure business relationships, predicting the conditions under which it is more efficient to outsource activities to third parties rather than perform them internally (e.g., Coase, 1937; Williamson, 1985). From this perspective, shirking and other forms of opportunism are the primary threats. This suggests that mechanisms other than independence—such as control—might motivate high-quality auditor performance.

Auditor independence

Research reveals how auditor independence is enhanced or eroded by the incentive structure of the relationships between clients and the third-party auditors they hire. Auditors tend to produce assessments biased in favor of their audited clients when the client pays for the audit (Duflo et al., 2013; Jiang, Stanford, and Xie, 2012), when auditors can cross-sell non-auditing services to the audited client (Causholli et al., 2014; Pierce and Toffel, 2013), or when auditors have long-term relationships with the audited entity (Griffin and Tang, 2011; Short, Toffel, and Hugill, 2016). Institutional conditions that can foster more independent auditor judgment include superior professional credentials (Short, Toffel, and Hugill, 2016), greater audit firm reputational pressure or legal pressure (DeAngelo, 1981; Francis and Wilson, 1988; Lytton and McAllister, 2014), and the gender composition of audit teams (Short, Toffel, and Hugill, 2016). Finally, research shows that cognitive biases, including bounded awareness, can limit auditors' ability to identify violations (Krogh, Roos, and Slocum, 1994; Kumar and Charkrabarti, 2012) and erode

their performance in repeated audits (Short, Toffel, and Hugill, 2016). Moreover, higher workloads and extended workdays lead to less-stringent inspections (Ibanez and Toffel, 2020).

Although this literature highlights important factors that shape auditor incentives and performance, it has several limitations. Because most studies have been conducted in the financial auditing context, the research has focused on third-party auditors serving firms that are *both* their clients and the entities they audit. Few studies address how factors shape incentives and auditor performance in contexts such as business partner monitoring in which the client and the audited entity are distinct firms. Moreover, second-party auditors have been largely ignored in the literature. This is a significant blind spot because it is not clear that second-party auditors are influenced by the same incentives as third-party auditors. Consequently, existing studies provide little guidance for the choice many firms face of whether to use second- or third-party auditors to monitor their business partners. The few studies that examine companies' choice of auditors tend to focus on how the third-party auditor selection process can be designed to maximize auditor independence and performance (e.g., Dao, Raghunandan, and Rama, 2012; Mayhew and Pike, 2004) rather than on the anterior question of whether auditing functions should be outsourced to third parties or performed in-house. Third, the literature's focus on how different types and levels of independence incentivize auditor performance leads it to overlook the possible influence of other incentives. Our study looks to research on outsourcing to theorize alternative mechanisms to incentivize high-quality auditor performance.

Outsourcing and supplier opportunism

The TCE perspective

To identify these mechanisms, we look to studies of firm boundaries and particularly to the transaction cost economics (TCE) literature, which focuses on firms' decisions about when to

make or buy a particular product or service. A foundational puzzle animating this literature is how firms should govern transactions in light of a key tradeoff they face when relying on external versus internal production. TCE assumes that outsourced suppliers can produce goods and services more cost-efficiently, such as by pooling demand from multiple buyers to achieve better economies of scale and more quickly climb learning curves. However, engaging with outsourced suppliers poses a risk: they might engage in quality shirking, a form of opportunistic behavior in which a supplier "is inclined to deliberately underperform" (Handley and Benton, 2012: 58) to earn "an immediate benefit in the form of a cost saving" (Wathne and Heide, 2000: 41).

Control is the key mechanism firms use to mitigate shirking by suppliers. Firms can use contracting and monitoring to mitigate supplier opportunism, but effective monitoring can be thwarted by high costs and asymmetric information, especially when supplier quality is not readily apparent in their products (Parmigiani, 2007: 290) or services (Jia, 2018). Firms that insource have much more control over production, not only because monitoring is easier but also because the organizational incentives of the goods' producers and users can be more directly aligned. However, firms that insource production might forfeit the efficiencies of outsourcing.

Our work builds on TCE in three ways. First, TCE predicts that a firm will govern each transaction by either purchasing from outsourced suppliers or producing the good or service in-house, depending on each transaction's risk of opportunism. TCE is particularly concerned with opportunism that arises when transaction-specific assets are required, such as specialized capital, and how easily a buying firm can manage such opportunism (e.g., via contractual terms and monitoring). TCE therefore predicts one governance mode per transaction—make *or* buy—but does not address why some firms might both make *and* buy, as occurs in the supply chain

auditing context we study. Second, TCE theory highlights ease of monitoring as a key influence on the make-or-buy decision, but it is silent on the make-or-buy decision for the monitoring function itself—the issue we study. Third, empirical work based on TCE has focused on predicting or validating transaction governance choices rather than evaluating the quality of supplier performance associated with those choices (Jia, 2018; Macher and Richman, 2008).

Concurrent sourcing

The literature on concurrent sourcing theorizes that firms gain several advantages by both making *and* buying the same good or service.[1] That literature has largely sought to identify factors that encourage concurrent sourcing and most of the empirical work has relied on perceptual measures collected in surveys. For example, survey evidence indicates that firms in the machinery and transportation equipment sectors that report greater difficulty monitoring the performance of outsourced suppliers are more likely to engage in concurrent sourcing (Heide, 2003). Another survey-based study revealed that concurrent sourcing was more likely when buying firms reported moderate levels of factors that have traditionally distinguished make-versus-buy decisions, such as how difficult it is to measure the quality of the supplied good (Parmigiani, 2007). Parmigiani and Mitchell (2009) found that concurrent sourcing of complementary goods is more likely when buying firms perceive that both they and external suppliers have the relevant skills and knowledge to produce them. A handful of studies have gone beyond predicting when concurrent sourcing is more likely to predicting its optimal levels (e.g., Puranam et al., 2013; Sako et al., 2016).

We extend this literature by examining how concurrent sourcing affects third-party

[1] Our focus on make-and-buy is one portion of the concurrent sourcing literature that also refers to this concept as plural sourcing (Jia, 2018; Puranam et al., 2013), taper integration (Harrigan, 1986), plural form (Bradach, 1997; Bradach and Eccles, 1989; Dutta et al., 1995), and plural governance (Heide, 2003).

suppliers' performance, an issue that has been theorized (as described below) but not empirically investigated. Relying on a panel dataset of observations from which we can observe actual performance, we identify how changes in concurrent sourcing subsequently affect suppliers' performance. This enables us to overcome two key limitation of the survey-based studies that have dominated the empirical work in this literature: their cross-sectional nature that reveals contemporaneous correlations, and their reliance on buyers' perceptions rather than more objective performance measures.

HYPOTHESES

To better understand how independence and control are associated with auditor performance, we examine relationships between buyers, suppliers, and auditors in a supply chain auditing context featuring two potential layers of opportunism and uncertainty. First, buyers use supply chain auditing as a governance mechanism to reduce uncertainty about supplier practices and mitigate supplier opportunism in settings in which suppliers might seek to reduce short-term operating costs by shirking on adherence to buyer-imposed labor standards. Buyers often outsource this monitoring function to third-party auditors, but doing so creates a second layer of uncertainty and potential opportunism. The quality of third-party audits features high uncertainty because audits are opaque and entail highly discretionary effort, creating ample room for third-party auditors to shirk their responsibilities or shill for the audited supplier. While these conditions would seem to pose a sourcing dilemma for buyers, most buyers nonetheless outsource auditing, bowing to the assumption that auditor independence alone will ensure that outsourced audits are conducted rigorously and reliably. Our hypotheses examine this assumption by theorizing whether independence or alternative mechanisms produce higher-quality auditor performance.

Independence through outsourcing

Auditor independence has long been the touchstone of audit quality across a wide range of monitoring contexts (Ammenberg, Wik, and Hjelm, 2001; Levitt, 2000; McAllister, 2012). Formal independence from the audited party is thought to attenuate auditor bias and avoid their shilling for audited entities by reducing a key conflict of interest that the auditor would face if performing a first-party audit of their its own organization (Kouakou, Boiral, and Gendron, 2013). The conflict of interest arises from the fact that first-party auditors stand to gain from their employer's malfeasance and thus have incentives to overlook it (Kraakman, 1986). This conflict of interest is also theorized to compromise the quality of second-party auditing in such contexts as supply chains, in which firms employ in-house auditors to monitor suppliers (Lebaron and Lister, 2015). These second-party auditors might overlook suppliers' opportunistic behavior because their employer stands to benefit if it provides competitive advantages, such as lower-cost and thus lower-priced supplies. Conversely, third-party auditors do not directly benefit from suppliers' opportunistic behavior, so they have less incentive to overlook it (Coffee, 2004).

In addition to theorized adverse incentives that undermine first-party or second-party auditor performance, some have theorized that reputational concerns incentivize high-quality performance by third-party auditors. This theory holds that third parties who monitor a variety of clients will not jeopardize their reputation in the broader market for the potential gains from performing a low-quality audit for one client (DeAngelo, 1981; Kraakman, 1986). Some studies show that reputation preservation can indeed reduce the likelihood that monitors succumb to client pressure to hide wrongdoing (Francis and Wilson, 1988; Ljungqvist et al., 2007). This reputation preservation incentive should apply to supply chain auditing contexts because third-party firms, which audit on behalf of many brands, are likely to use their reputations for high-

quality audits to attract and retain clients.

We therefore hypothesize:

Hypothesis 1a: Third-party auditors outperform second-party auditors.

Insourcing audits to increase control

Insights from the TCE literature, however, suggest the opposite relationship: in some settings, a buyer's own employees—such as second-party auditors—have stronger incentives than outsourced suppliers—such as third-party auditors—to exhibit high-quality performance. First, as discussed above, third-party auditors have incentives to shirk their duties to reduce their costs. Second, third-party auditors are not well positioned to account for all potential externalities faced by their clients, particularly those arising in high–reputation-risk contexts. For example, one audit conducted hastily (to save the auditor time and cost) that overlooks fire safety violations could result in many worker injuries and deaths, with devastating business and reputational consequences for the MNC buyer. However, such a tragic outcome would have much less impact on the third-party auditor, because NGOs and the media tend to hold buyers, not third-party auditors, accountable for ensuring adequate working conditions in global supply chains.

Studies in the TCE literature demonstrate that firms choose to insource transactions that ordinarily would be outsourced when those transactions expose the firm to externalities in the form of reputational spillovers (Mayer, 2006; Nickerson and Silverman, 2003). Nickerson and Silverman (2003), for example, argue that trucking companies insource to their own drivers certain hauls—those for which poor performance could seriously imperil the trucking company's reputation or broader operations—because the companies can better calibrate incentives to control the behavior of their own employees than that of third-party drivers. We argue that a

firm's superior control of its own employees can be an alternative to independence as a mechanism to produce high-quality auditor performance. We hypothesize that second-party auditors are thus likely to conduct higher-quality audits than third-party auditors when the risk of reputational spillovers is significant, as it is in the context of suppliers' labor standards compliance.

We therefore propose the following alternative hypothesis:

Hypothesis 1b: Second-party auditors outperform third-party auditors

Control through concurrent sourcing

Another factor that can affect the quality of third-party auditors' performance is whether buyers engage in concurrent sourcing by simultaneously using second- and third-party auditors. Concurrent sourcing bolsters firms' ability to control their suppliers through three different mechanisms. First, the buyer firm can better monitor outsourced production. Several studies argue that concurrent sourcing enables buyer firms to learn from their in-house suppliers about production costs, technologies, and metrics, all of which enhances buyer firms' ability to negotiate with and monitor outsourced suppliers (Bradach and Eccles, 1989; Heide, 2003; Jia, 2018; Puranam et al., 2013). Moreover, concurrent sourcing has been theorized to help firms evaluate and compare the performance of third-party suppliers to that of internal suppliers. For instance, concurrent-sourcing firms can tap their in-house production knowledge to better assess whether the poor quality of products furnished by a third-party supplier is due "to genuine manufacturing problems or to surreptitious supplier cheating" (Heide et al., 2014: 1166). In addition, firms can benchmark performance between their in-house and third-party suppliers to "ratchet" up performance by creating "a virtuous cycle of continuous improvement" (Puranam et al., 2013: 1151).

Second, in the supply chain auditing context, concurrent sourcing physically facilitates "monitoring the monitors" by placing in-house auditors on the ground in the same markets as third-party auditors. This makes it easier for in-house auditors to conduct spot-checks of third-party auditors' work by visiting the same factories. Even if such spot-checking is done sparingly to minimize duplicative auditing cost, the market presence of second-party auditors signals to third-party auditors that the MNC client can readily assess their work, which should deter opportunism and encourage higher-quality performance.

Third, concurrent sourcing provides a plausible threat that the buyer firm can readily increase insourcing if dissatisfied with third-party auditors. Competition can thus deter opportunism and bolster performance (Dutta et al., 1995; Heide et al., 2014; Puranam et al., 2013).

These theories, largely developed with an eye toward the procurement of industrial products, can also be usefully applied to service contexts such as auditing. Service quality is often particularly difficult to monitor, due in part to the intangibility of services (Alzaydi et al., 2018). Service quality assessment is further complicated in an auditing context because the auditors perform their work at the client's suppliers and not at the client's own establishments, making the service provision far more difficult and costly for the clients themselves to monitor.

We theorize that MNCs introducing second-party auditors into a market through a concurrent sourcing strategy will discourage their third-party auditors from engaging in opportunism. We therefore hypothesize:

Hypothesis 2: Concurrent sourcing improves the performance of third-party auditors.

Independence erosion from repeated encounters

Audit quality is likely to depend not only on the relationship between auditors and their

branded clients, but also on relationships between auditors—both third-party and second-party—and the suppliers they monitor. A substantial body of research demonstrates that ongoing relationships between individual auditors and the firms they audit can create biases that dampen auditor performance (Jin and Lee, 2018; Lennox, Wu, and Zhang, 2014; Short, Toffel, and Hugill, 2016); some regulators have mandated periodic rotation of individual auditors. For example, the US Sarbanes-Oxley Act and Securities and Exchange Commission regulations require that the two lead auditor partners of a third-party audit firm be rotated off a client engagement after five years, followed by a five-year "time-out" period from working with that client (Securities and Exchange Commission, 2003). Less is known, however, about the implications of the repeated use of the same audit firm, despite some calls for audit firm rotation to avoid the erosion of auditor performance.

It has been widely observed that individual auditors are less stringent when auditing entities with which they have ongoing relationships. By building a relationship with audited suppliers, auditors risk losing the emotional distance needed to be an effective monitor. Compromised emotional distance may lead individual auditors to become sympathetic to a supplier's position (Moore et al., 2006; Short, Toffel, and Hugill, 2016) or render them more vulnerable to bribery (Khalil and Lawaree, 2006; Montiel et al., 2012). In addition, repeated interactions between individual auditors and the suppliers they audit can exacerbate the problem of individual auditors' "bounded awareness"—the suite of cognitive mechanisms that lead individuals to focus on certain, limited categories of information (Chugh and Bazerman, 2007). Consistent with this, one study found that when individual auditors returned to the same supplier, they recorded fewer violations than auditors who were new to the factory (Short, Toffel, and Hugill, 2016).

In addition to these *individual* emotional and cognitive factors shaping auditor performance, auditors are also affected by the *organizational* environment of the auditing firm in which they are embedded. Specifically, an audit firm's organizational structures, culture, and routines (to which we refer collectively as "organizational structures") are likely to shape its employees' cognitive heuristics via mechanisms such as employee training (Metzger et al., 1993), organizational norms (Trevino, Butterfield, and McCabe, 1998), codes of conduct (Cressey and Moore, 1983), and social networks (Brass et al., 1998). Thus, an auditor's performance is likely to decline not only when the auditor has already audited a given entity, but also when coworkers have audited it. For example, auditors relying on organizational structures conveyed by their firm might, in consecutive factory visits, focus on similar violation types, over-search areas of the factory previously covered by their colleagues (where issues are more likely to have been resolved in response to prior audit findings), or under-search other areas that lead them overlook a common set of violations that were not stressed in firm training materials.

We therefore hypothesize:

Hypothesis 3: Auditor performance is higher when auditing entities whose most recent audit was conducted by a different auditing company.

DATA AND MEASURES

Empirical context and measures

We obtained factory audit data from Primark, a large retailer of apparel, home goods, and beauty products with 384 stores in thirteen countries throughout Europe and the United States. Primark, known for selling trendy fashion items at low prices, monitors its supply chain factories to assess their compliance with its supplier code of conduct whose terms are closely aligned with the Ethical Trading Initiative's base code of labor practices. These audits are conducted by second-party auditors (its employees) or external third-party auditors, who review documents, tour the

factory to observe conditions, and interview workers.[2] The MNC selects the auditors and pays for all audits, and the vast majority of both second- and third-party auditors are local to the supplier factory's country or region.[3] Factories are then assigned an overall "social audit rating" (Red-Critical, Red, Amber, or Green) based on the number and severity of the violations.

We were provided data on all 15,935 audits conducted for Primark during 2007–2017 by second- and third-party auditors. The factories were in 52 countries, but most audits were conducted in China (50% of our sample), India (16%), and Bangladesh (9%).

Dependent variable

We measure factory compliance with the MNC buyer's code of conduct as the number of *new violations* recorded during an audit. Prior research in similar compliance contexts, such as government health and safety regulation (Braithwaite and Makkai, 1991; Gray and Shadbegian, 2005), food inspection standards (Ibanez and Toffel, 2020), and private codes of conduct (Ang et al., 2012; Oka, 2010; Short, Toffel, and Hugill, 2016), uses violation counts recorded in inspections to measure compliance. Whereas these studies used total violation counts, including newly recorded violations and any unresolved violations identified in prior audits, we use only the newly recorded violations in order to isolate the focal auditor's ability to identify new issues and to avoid the potential confound of a factory's unobserved heterogeneous ability or willingness to resolve previously identified violations.

The code of conduct contains about 90 binary compliance elements that span 12 violation categories that track the Ethical Trading Initiative's Base Code, which is based on International Labour Organization (ILO) core conventions: child labor, communication of the code of conduct

[2] 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.
[3] Primark selects third-party auditors based on their capabilities, capacity, geographic coverage, skills profile of auditors, and other factors.

to workers, discrimination, harsh worker treatment, environmental requirements, freedom of association, legal requirements, living wage, regular employment, working conditions, working hours, and voluntary employment.

Independent variables

To measure the level of independence and control between the MNC buyer and the auditor, we coded *second-party auditor* as 1 when an audit was conducted by a second-party auditor (the MNC's employees) and as 0 when conducted by a third-party auditor.

We identified factory audits occurring in concurrent sourcing markets by coding *concurrent source* as 1 when second-party auditors conducted between 10 percent and 90 percent of all audits in a particular country-year, and 0 otherwise.[4] This 10-percent-to-90-percent range is consistent with other empirical studies of concurrent sourcing (e.g., Parmigiani, 2007).[5]

We constructed three binary variables to measure the independence between the audited factory and the audit firm. *Different audit firm* refers to a factory audit conducted by a different audit firm (and thus a different lead auditor) than the factory's prior audit was. *Different lead auditor (same audit firm)* refers to an audit conducted by a different lead auditor but from the same audit firm as the factory's prior audit. *Same lead auditor (same audit firm)* refers to an audit conducted by the same lead auditor and audit firm as the factory's prior audit was. Collectively, these variables enable us to measure the effect of audit firm rotation while accounting for lead auditor rotation, which prior research demonstrates can affect the number of violations cited (Short, Toffel, and Hugill, 2016).

Control variables

[4] That is, *concurrent source* equals 0 when either second-party auditors or third-party auditors conducted nearly all (more than 90 percent) of the audits in a country-year.

[5] Robustness tests that measured concurrent sourcing using several alternative ranges (1% to 99%, 5% to 95%, 15% to 85%, 20% to 80%, and 25% to 75%) produced nearly identical results, as discussed below.

We created a number of audit-level variables to control for factors that might affect auditor performance. *Unannounced audit* equals 1 when the audited factory had no advanced notice of the audit, and 0 otherwise. Factories receiving advanced notice have more time to resolve or conceal potential new violations, which might reduce the number of violations recorded.

Four binary variables differentiate audit scope, which could affect the number of violations detected. *New factory audit* refers to a new supplier factory audit, which is conducted before the MNC buyer has placed an order. *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.[6] *Re-audit* refers to a routine audit conducted roughly annually after the factory's prior new factory audit or re-audit. *Follow-up audit* refers to an audit typically conducted a few months after a factory's previous audit yielded a "Red" or "Red-Critical" overall social audit rating. We also created versions of these variables that correspond to the factory's prior audit and refer to them as "prior audit" variables (e.g., *re-audit (prior audit)*).

Because prior research has found that auditor gender can influence performance (Short, Toffel, and Hugill, 2016), we created 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.[7] We used genderize.io, which uses profile information across major social networks to predict gender and has been used in other scholarship for this purpose (e.g., Greenberg and Mollick, 2017; Lundberg and Stearns, 2019; Rubineau and Fernandez, 2015).

We also measured factory-level characteristics that might influence auditor performance. Because factories that have worked longer with the MNC buyer might have resolved more code-

[6] For example, auditors may need to return to a new supplier factory if that factory does not have all documentation prepared for the auditors on their first visit.
[7] Our data includes the names of lead auditors but not of other audit team members and does not include gender.

of-conduct issues, we created *factory tenure* as the number of years the MNC buyer had worked with the factory at the time of the focal audit. Because the proportion of a factory's output supplied to the MNC buyer at the time of the focal audit might influence how beholden it is to that buyer—and perhaps how motivated it is to adhere to the code of conduct—we created *percent supplied to MNC buyer* as the proportion of factory output supplied to the MNC buyer at the time of the focal audit. Because factories with more employees might have more violations, we create *total workers (log)* as the number of factory workers at the time of the audit.

Table 1 provides summary statistics and Table A-1 in the Appendix reports correlations.

EMPIRICAL ANALYSIS

Empirical model

We estimate the effect of our hypothesized variables on auditor performance with this model:

$$Y_{id} = F\left(\beta_1 X_{id} + \beta_2 \lambda_{id} + \beta_3 \rho_{i,d-1} + \beta_4 \alpha_i + \beta_5 \delta_y + \beta_6 \gamma_s + \epsilon_{id}\right) \quad (1)$$

Y_{id}, is our measure of auditor performance, the *new violations* recorded in the audit of factory *i* in audit *d*. *F()* represents the Poisson function, X_{id} represents our hypothesized explanatory variables: *second-party auditor* (H1), *concurrent source* (H2), and *different audit firm* (H3). λ_{id} represents the control variables listed above (*unannounced audit*, *same lead auditor (same audit firm), limited scope audit*, *follow-up audit, re-audit, female lead auditor, factory tenure, percent supplied to MNC buyer,* and *total workers (log)).* $\rho_{i,d-1}$ refers to factors associated with the factory's prior audit that we control for; namely *limited scope audit (prior audit)*, *follow-up audit (prior audit),* and *re-audit (prior audit). Different lead auditor (same audit firm)* serves as the omitted category for *different audit firm* and *same lead auditor (same audit firm).* The focal (and prior) audit versions of *new factory audit* serve as the omitted category for the focal (and prior) audit versions of *limited scope audit*, *re-audit* and *follow-up audit*.

α_i refers to supplier-factory fixed effects, which control for the factory's industry, location, and other time-invariant factory-specific effects. δ_y, refers to audit-year fixed effects, which control for secular changes that might affect factory violations or auditor stringency. We also control for audit sequence—whether the audit is the factory's first audit, second, and so on—by creating a dummy variable for each audit sequence value (γ_s) 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, Short, and Toffel, 2019).

Identification and results

We estimate the models that test our hypotheses using quasi-maximum likelihood estimation of the Poisson likelihood function, which is consistent even for dependent variables that are not Poisson-distributed (Cameron and Trivedi, 1998).[8] We report results in Table 2.

Comparing second- and third-party auditors' performance

For Equation 1 to yield an unbiased estimate of the effect of an audit being conducted by a *second-party auditor* (versus a third-party auditor) on *new violations* (Y_{id}), no unobserved variables should be correlated with both Y_{id} and the decision to assign a second- versus third-party auditor. We questioned whether this assumption held in our data for a few reasons. First, our interviews with Primark employees revealed that it prioritizes assigning second-party auditors to conduct new factory audits. Second, exploratory analysis revealed that (a) second-party auditors were more likely to be assigned to audit factories that had worse prior audit results[9] and (b) factories' prior and focal audit results were correlated. It is therefore possible that

[8] This estimation methodology also avoids the incidental parameter problem exhibited by regression models suited for count data with many fixed effects, such as a negative binomial model with fixed effects.
[9] This finding is consistent with the TCE literature's prediction that firms internalize transactions that can impose larger externalities on the rest of the firm's operations in the form of reputational spillovers (Mayer, 2006; Nickerson and Silverman, 2003).

second-party auditors were more likely to be assigned to factories that had more violations, which confounds our ability to compare the performance of second- and third-party auditors.

To address this concern, we developed a matched sample of audit pairs—one audit assigned to a second-party auditor, the other to a third-party auditor—that were otherwise as similar as possible in their prior audit results and several other factors that might be correlated with both auditor assignment and our outcome variable.[10] Specifically, we used coarsened exact matching[11] to exactly match on: (1) a coarsened version of *new violations* recorded in the factory's prior audit; (2) the factory's prior audit's *social audit rating*; (3) *unannounced audit*; (4) audit scope (*new factory audit*, *limited scope audit*, *re-audit*, or *follow-up audit*); (5) *concurrent source,* (6) the factory's country, (7) *audit year*, and (8) *audit sequence* (top-coded at the 95th percentile, the eighth audit).[12] The resulting matched sample includes 3,526 audits of 2,230 factories.[13] We assessed balance across 26 variables: 5 predictors of assignment to address endogeneity concerns and 21 predictors of performance to ensure common support across second- and third-party audits. Comparing group means between second- and third-party audits in this matched sample indicates that 22 of the 26 variables are balanced (meaning that the group means were statistically indistinguishable at the 10% level), a substantial improvement over the

[10] One potential solution could be to include the lagged number of new violations as a control variable. However including this (lagged dependent variable) as a regressor would bias our estimates due to a dynamic panel problem. The typical solution of using dynamic panel estimators requires second and third lagged dependent variables as instruments and is therefore infeasible in our context of short panels (e.g., only 28% of suppliers in our dataset had at least four audits).

[11] We rely on the *cem* command in Stata to implement our matching procedure, which implements the method articulated in Iacus, King, and Porro (2008).

[12] We include both a supplier factory's prior audit social audit rating and its number of new violations as measures of prior auditor performance. While an audit's social audit rating correlated with the number of violations recorded in the audit, the social audit rating also accounts for violation severity. Because both measures are missing for a factory's first audit, we recoded those instances to 0 and include in our models dummy variables coded 1 to designate those instances.

[13] For summary statistics of this matched sample, see Table A-2.

mere two that were balanced in the full sample.[14,15] We controlled for these unbalanced variables in our regression model and in a robustness model.[16]

We used this matched sample to test H1 and report the results as Model 1 of Table 2, clustering standard errors by factory. The positive coefficient for *second-party auditor* (β=0.27; $p<0.01$; IRR=1.30) reveals that second-party auditors, on average, record 30 percent more new violations than third-party auditors and supports Hypothesis 1b, which predicted that second-party auditors outperform third-party auditors. The average marginal effect of 1.9 indicates that while third-party audits in the H1 sample record an average of 6.5 new violations, second-party auditors record an average of 8.4.

We conducted several robustness tests to explore the sensitivity of our main H1 results. First, we assessed the extent to which our results were driven by the matched data sample on which they were estimated by re-estimating Model 1 on the full sample of audits. Second, to assess whether our results are sensitive to how performance is measured, we re-estimated Model 1 using three alternative dependent variables: *new violations top-coded at the 95th percentile* (19 new violations) to compress outliers, *major new violations* to assess whether results hold even after omitting minor violations, and *total violations,* a common metric in the literature. These four models yield coefficients on our hypothesized variables that are very similar in magnitude

[14] Table A-3 reports the covariate balance of the matched sample and compares it to that of the full sample. p-values were obtained from t-tests for non-binary variables and proportional tests for binary variables. The average standardized bias, which measures balance in terms of the similarity in covariate distributions, is 0.6 in the matched sample, a substantial improvement over its -3.1 value in the unmatched sample. Standardized bias is calculated as the difference in group means divided by the average standard deviation. Values closer to zero imply greater similarity in distributions.

[15] The kernel density graphs in Figure A-1 in the Appendix show that in the matched sample, the distribution of the variables for second-party audits and for third-party audits is very similar.

[16] Our primary models control for three of these four variables that remained unbalanced: *different audit firm*, *different lead auditor (same audit firm),* and *same lead auditor (same audit firm).* A robustness test model that includes the fourth unbalanced variable, *second-party auditor (prior audit),* yields nearly identical results (not reported).

and direction to our primary results,[17] indicating that our H1 results are robust to this alternative sample and to these metrics.

Assessing how third-party auditors' performance differs with concurrent sourcing

When testing H2, we restrict our sample to audits conducted by third party auditors.[18] To isolate the effect of transitioning from a third-party market to a concurrent sourcing market, we exclude the few instances (75 audits in 7 country-years) in which the reverse transition occurred—a supplier factory transitioned from a concurrent sourcing market to a third-party market.[19] Because Equation 1 includes factory-level fixed effects, the effect of concurrent sourcing on new violations for third-party auditors is identified when factories switched from receiving third-party audits in a predominantly third-party–audited market (in which second-party auditors conducted less than 10% of audits in a country-year) to a concurrent sourcing market (in which second-party auditors conducted between 10% and 90% of audits in a country-year).[20] To identify the estimation sample, we first consider all 10,071 third-party audits for which we had data on all measures in Equation 1.

Model 2 in Table 3, where we cluster standard errors by country because *concurrent source* is measured at the country-year level, reports the results of our testing H2. The positive coefficient on *concurrent source* ($\beta = 0.26$; $p<0.01$; IRR=1.30) indicates that third-party auditors record an average of 30 percent more new violations per audit after the MNC has deployed concurrent sourcing in the factory's market, which supports H2. The average marginal effect

[17] Results are reported in Table A-4.
[18] Summary statistics of this sample are reported in Table A-2.
[19] Including these audits does not materially change our results.
[20] To avoid endogeneity, our identifying assumption is that this market transition was not sparked by unobserved factors that might have both affected the decision to concurrently source and independently affected third-party auditor performance. Interviews with the MNC reveal that its decision to have its second-party auditors enter a particular country or to scale up its auditing in particular countries was based on growing order volume procured from the suppliers in that country, which is plausibly unrelated to the quality of third-party auditor performance.

indicates that third-party audits in concurrent source markets average 5.5 new violations, which is 1.2 more than the 4.3 average in markets served exclusively or nearly exclusively by third-party auditors.

We found similar results when, as robustness tests, we identified concurrent sourcing markets using five alternative thresholds (1% to 99%, 5% to 95%, 15% to 85%, 20% to 80%, and 25% to 75%) instead of our primary 10% to 90% approach (results reported in Table A-5). Moreover, we find similar results when we re-estimate Model 3 using the three alternative dependent variables listed above (see Table A-6). Collectively, these results indicate that our primary results are robust to alternative thresholds when measuring our independent variable and to alternative measurements of the dependent variable.

Assessing how auditors' performance differs with audit firm rotation

To obtain an unbiased estimate of the effect of audit firm rotation on auditor performance (H3), the decision to assign a different auditing firm to a supplier factory's focal audit should be uncorrelated with the factory's prior audit results. Interviews with the MNC reveal that the assignment of audit firms to factory audits (and thus rotation) is based on auditor availability, which is unlikely to be correlated with prior audit results. Thus, we test H3 on the entire sample of audits for which we had data on all measures described above.

The results of Model 3, which tests H3, are reported in Column 3 of Table 3, where we cluster standard errors by factory. The coefficient on *different audit firm* ($\beta = 0.24$; $p < 0.01$; IRR = 1.27) indicates that audits with rotated audit firms report 27 percent more new violations than those with rotated lead auditors within the same firm (the omitted category). The average marginal effect indicates that audits conducted by firms other than that which conducted the factory's prior audit report an average of 1.5 more new violations than the 4.0 average for audits

conducted by the same firm as last time but with a different lead auditor. In supplemental analyses, we found that this result held whether the focal auditors were second-party auditors or third-party auditors. [21]

To assess whether the audit firm assignment process might be driving our H3 results, we re-estimated Model 3 on two alternative subsamples. First, we omitted from the sample those audits for which intentional audit firm assignment might have been especially tempting due to an unusually pronounced decline in *new violations*.[22] Second, we used coarsened exact matching to create a matched sample that matched one audit conducted by an audit firm that had not conducted the factory's prior audit (audit firm rotation) to an audit conducted by the audit firm that had conducted the prior audit (not rotation).[23] The results of re-estimating Model 3 on these two subsamples, reported in Columns 1 and 2 of Table A-7, are quite similar to those in our primary model, suggesting that our primary estimates of H3 are robust to plausible scenarios that seek to further reduce potential endogenous audit firm rotation. Re-estimating Model 3 to instead predict each of the three alternative dependent variables described earlier provides evidence that the relationship predicted by H3 is robust across of these alternative metrics, as indicated in Columns 3–5 of Table A-7.

As an extension, we explore the extent to which audit firm familiarity might attenuate the

[21] Specifically, we re-estimated Model 3 on the subsample of second-party focal audits, and separately on the subsample of third-party focal audits. The results of both models (not reported) yielded a positive coefficient on *different audit firm* that was statistically significant at the 5% level.

[22] Specifically, we created an improvement score by calculating the percent reduction in the number new violations, comparing a factory's prior audit to its preceding two audits (or single preceding audit when two were not available). Of the 7,751 audits with non-missing improvement scores, we omitted the 1,948 whose percent declines were the highest 25th percentile (that is, when improvement was 80% or greater). We supplemented the remaining 5,803 audits with the 8,184 audits for which we could not calculate improvement—when the focal audit was the factory's first or second audit or when the previous two audits averaged 0 new violations—yielding a subsample of 13,987 audits of 4,685 factories.

[23] Specifically, we matched on a coarsened version of new violations (prior audit), social audit rating (prior audit), second-party auditor (versus third-party auditor), audit scope (new factory audit, limited scope audit, re-audit, and follow-up audit), supplier country, audit sequence (top-coded at 95th percentile), and audit year. The resulting matched sample contains 6,046 audits of 2,231 factories.

benefits of audit firm rotation, compared to rotating to an audit firm that has never audited the factory before. To do so, we decomposed *different audit firm* into *new different audit firm*—which equals 1 for an audit by a firm that has never audited the factory before, and 0 otherwise—and *familiar different audit firm*—which equals 1 when the audit firm has conducted audits at the factory before but not the most recent one.[24] In the results reported in Table A-8, the average marginal effects indicate that, compared to audits conducted by the same firm as last time but with a different lead auditor (the omitted category), audits rotated to a *new different audit firm* report an average of 1.8 more new violations, and audits rotated to a *familiar different audit firm* report an average of 1.0 more. A chi-squared test comparing the corresponding coefficients confirms that this 0.8 difference is statistically significant at the 1-percent level ($\chi^2 = 15.24$; $p < 0.01$).

Control variables

The coefficients on several controls in the models reported in Table 2 warrant discussion. The negative coefficient on *unannounced audit* across all specifications indicates that audits for which factories are given no advance notice yield more new violations than pre-announced audits (average marginal effects ranging from 0.8 to 2.8 new violations). This result is directionally in line with prior findings that audits conducted with less advance notice report more violations (LeBaron and Lister, 2015; Short, Toffel, and Hugill, 2016).

The coefficients on *female lead auditor* in Models 2 and 3 indicate that female-led audits report an average of 0.8 to 0.9 more new violations ($p < 0.01$) than male-led audits. The coefficients on *same lead auditor (same audit firm)* across the three models indicate that audits conducted by the same lead auditor (at the same firm) as was the factory's prior audit report 1.3

[24] We estimate this model on all audits and cluster standard errors by factory.

to 2.3 fewer new violations ($p < 0.01$) than audits for which the lead auditor is rotated but not the audit firm. Both results are consistent with the findings by Short, Toffel, and Hugill (2016) that more *total* violations were reported in audits conducted by audit teams that had at least one female auditor and that had not visited the factory before.

DISCUSSION

Our study of auditing in the context of multi-layered supply chain assurance regimes reveals that auditor performance is shaped not only by the potential bias and conflicts of interest traditionally addressed through independence criteria, but also by the incentives to behave opportunistically that all suppliers of goods and services have. Our approach allows us to identify novel audit sourcing arrangements—rotating audit firms and controlling auditors through insourcing or concurrent sourcing—that can promote better auditor performance. Our work makes five significant contributions to the literatures on auditor independence and outsourcing decisions and has important managerial implications, all of which we discuss below.

First, by foregrounding the "make-or-buy" decision underlying the choice of auditors, we identify cross-cutting issues that can affect auditor performance and we generate new insights about how to incentivize better auditor performance. While the focal problem in the auditor independence literature is auditors *shilling* for their clients, TCE theory suggests that firms should also worry about opportunism manifest as auditor *shirking*. Independence is not the solution to auditor shirking and, indeed, might exacerbate it. We show that in some circumstances, firms elicit better performance from auditors by controlling them rather than keeping them at a distance. To be clear, we do not claim that second-party auditors always perform better than third-party auditors. As we argue above, control is likely to be particularly important for firms facing high risk of reputational spillover from suppliers' mistakes—as

branded MNC buyers and many other firms do. And, of course, second-party auditors are likely to perform well only if their employer is committed to learning the truth about misconduct by its business partners. More research is needed to identify and test these and other circumstances that might differentially affect the performance of second- versus third-party auditors and other suppliers.

Second, our study provides a novel window into the role of second-party auditors in assurance regimes. Second-party auditors have been an afterthought in a literature focused almost exclusively on first- and third-party auditors. This is likely because financial auditing provides the model for designing audit regimes (Power, 1997) and, in that context, firms are legally required to outsource audits to third parties rather than perform first-party audits of their own financials. But even scholars and advocates focused on more complex assurance regimes involving multi-layered business relationships tend to assume that second-party auditors suffer from the same biases and conflicts of interest as first-party auditors (e.g., Lebaron and Lister, 2015; McAllister, 2012). To our knowledge, our study is the first to empirically test this assumption and we find it to be mistaken—or at least incomplete. Our findings suggest that knowledge established in studies of third-party auditors should not be applied automatically to second-party auditors. Rather, it is necessary to theorize and test how incentives might shape second- and third-party auditor performance differently in particular settings.

Third, and more broadly, our findings suggest the need to revisit conceptions of legitimacy in auditing. Many have observed that the function of auditor independence is to confer legitimacy on the auditing process and its results (Boiral and Gendron, 2011; Kouakou, Boiral, Gendron, 2013; Markell and Glicksman, 2014). Independence is a particularly important signal of legitimacy for outside stakeholders who have little other information by which to judge

audit quality (Darnall and Vazquez-Brust, 2018; Jia, 2018). But it is not clear that this visible symbol of audit legitimacy accurately reflects audit quality. Critical scholarship in the auditing literature suggests that there is a tension between auditing as a tool for information gathering and auditing as an accountability mechanism—or between reliability and external perceptions of legitimacy (e.g., Power, 1997). Our finding that second-party auditors outperform third-party auditors in our empirical context lends some support to critics' charge that third-party audits are symbolic and performative tools that convey legitimacy to outside stakeholders, rather than rational tools that reduce monitoring costs and information asymmetry (Power, 1997). Scholars, advocates, and managers who seek to design high-quality assurance regimes should carefully consider what role independence can play in promoting reliable audits, what its limitations are, and what other tools are available.

Fourth, in addition to demonstrating how disparate incentives can generate divergent second- and third-party auditor performance, we show that both types of auditor are subject to structural biases that limit their performance when returning to an audited entity. Our finding that both insourced and outsourced auditors perform worse when they follow visits by colleagues from their own audit firm is an important contribution to the literature on audit firm rotation. Although there has been much research on audit firm rotation, it has several limitations that our study overcomes. First, many studies framed in terms of audit firm rotation do not, in fact, study it. Instead, they study the relationship between longer audit firm tenure and audit quality and extrapolate to form conclusions on the value of audit firm rotation (e.g., Al-Thuneibat et al., 2011; Carcello and Nagy, 2004; Tepalagul and Lin, 2015). Second, studies that do examine audit firm rotation do so in financial auditing contexts in which (a) such rotation is mandatory and almost always anticipated by audit firms, (b) audits are legally required, (c) there is legal liability

for poor-quality audits, and (d) auditors are auditing the performance of their clients and not their clients' business partners (e.g., Blouin et al., 2007; Dyck, Morse, and Zingales, 2017; Ruiz-Barbadillo et al., 2009; Wang and Tuttle, 2009). These financial-auditing–specific characteristics may create incentives that are not present in other auditing contexts. We clarify how firm rotation can promote better auditor performance in the context of supply chain monitoring.

Finally, our study contributes to the TCE literature and the literature on concurrent sourcing by evaluating the relative performance of insourced, outsourced, and concurrently sourced production. To our knowledge, ours is the first study to do so. Both literatures focus almost exclusively on explaining transaction governance choices rather than evaluating how these choices affect the quality of supplier performance. Thus, our study responds to repeated calls for more attention to the outcomes of outsourcing decisions (Jia, 2018; Macher and Richman, 2008). In addition, we foreground the choice of monitors as a central concern for TCE theory. While many studies in the TCE literature theorize monitoring as a means to mitigate supplier opportunism, few consider how the monitoring function is itself subject to supplier opportunism and warrants careful sourcing consideration. Moreover, the insight we provide into the choice of auditors extends both the TCE and concurrent sourcing literatures by demonstrating how their theoretical frameworks apply in the understudied context of outsourcing services (rather than outsourcing the production of tangible goods).

Managerial implications

Our results have several important implications for managers monitoring their business partners' performance. First, our study challenges the prevailing wisdom that third-party audits are always the most reliable form of assessment. While such audits may be perceived as more

legitimate by outside stakeholders, managers are likely to learn more from their own in-house audit team about what is happening on the ground at their suppliers. Managers should periodically evaluate the effectiveness of their monitoring programs and work closely with third-parties to close any performance gaps detected. Managers who elect to perform audits in-house 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 reasons for their audit sourcing decisions and the benefits of different sourcing arrangements. Second, our findings help managers mitigate the tradeoffs they face when second-party audits yield higher-quality information but are more expensive than third-party audits. We demonstrate that managers need not sacrifice the typical cost-savings benefits of outsourced audits to bolster auditor performance. Integrating a concurrent sourcing strategy into an outsourced audit program can significantly improve the quality of outsourced audits. This finding is important for managers who wish to capture the reliability benefits of second-party audits but lack the resources or auditor availability to insource all audits. Finally, managers should consider rotating the firms performing audits of their business partners to ensure that their auditors do not fall into the rut of organizational routines that may cause them to overlook important information.

Limitations and future research

We acknowledge several limitations to our study. First, our data come from a single firm that uses both second- and third-party auditors, which raises questions about the generalizability of our finding that second-party auditors outperform third-party auditors. It has proven difficult for apparel MNCs to converge on a common audit protocol and to share audit results with each other, so there are no datasets that include multiple buyers as well as their second- and third-party auditors. We use the best data available to test our hypotheses and we encourage others to

test the generalizability of our findings and variations in the conditions under which they hold.

Second, despite being able to match on all variables that we believe are important to auditor assignment and performance, our reliance on observational data means that non-random auditor assignment, including the endogenous choice of concurrent sourcing, may occur along unobserved dimensions that could bias our results. Our interviews with MNC audit schedulers have not suggested that unobserved variables affect auditor assignment (and performance), but the possibility exists nonetheless.

Third, the proprietary audit data we obtained do not include variables to enable us to control for auditor skill or professional credentials, which prior research has shown do affect audit results (Short, Toffel, and Hugill, 2016). While our MNC's second-party and third-party auditors received comparable training on the MNC's code-of-conduct requirements and while our direct observations of audits being conducted did not expose differences in audit skill, we cannot rule out the possibility that training and skill might have affected audit results in ways we cannot observe. We note that the fact that the MNC selects and pays all its third-party auditors, rather than allowing suppliers to do so, eliminates a significant potential source of auditor bias.

Finally, given the limitations of our data, we can only theorize—but not empirically test—the underlying mechanisms driving auditor performance. For instance, it is difficult to conclude whether the performance gaps we document are due to a lack of effort on the part of third-party auditors or a lack of access to information. It is possible that suppliers are less transparent with third-party auditors because information provided in one third-party audit risks spilling over to audits conducted on behalf of other brands. However, it is also possible suppliers are less transparent with second-party auditors because they have a clear link to sourcing decisions.

These limitations create opportunities for future studies. Future research should investigate what combinations of independence and control promote effective monitoring of other types of business partners, such as franchisees, distributors, and purchasing agents, and whether these mechanisms can improve the performance of other service suppliers besides auditing. Future research could drill down into the mechanisms we theorize; for instance, testing auditor performance 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 performance gaps between second- and third-party auditors.

CONCLUSION

In a world of complex, multi-layered business relationships and intricate business, regulatory, and social demands, firms must vigilantly monitor transaction partners. While monitor independence is often presumed to be the principal determinant of performance, we demonstrate that sourcing decisions, such as insourcing or concurrent sourcing, also play a role. We theorize these performance differences by synthesizing literatures on auditor independence and outsourcing decisions. Our findings contribute significantly to these literatures and provide important managerial insights for designing and implementing more effective business partner monitoring strategies.

REFERENCES

Al-Thuneibat, A.A., Al Issa, R.T.I., and Baker, R.A.A. 2011. Do audit tenure and firm size contribute to audit quality? *Managerial Auditing Journal* 26(4): 317-334.

Alzaydi, Z.M., Al-Hajla, A., Nguyen, B., and Jayawardhena, C. 2018. A review of service quality and service delivery. *Business Process Management Journal* 24(1): 295-328.

Ammenberg, J., Wik, G., and 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.

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

Bird, Y., Short, J.L. and Toffel, M.W. 2019. Coupling Labor Codes of Conduct and Supplier Labor Practices: The Role of Internal Structural Conditions. *Organization Science 30*(4): 847-867.
Blouin, J., Grein, B.M., and Rountree, B.R. 2007. An analysis of forced auditor change: The case of former Arthur Andersen clients. *The Accounting Review* 82(3): 621-650.
Boiral, O., and Gendron, Y. 2011. Sustainable development and certification practices: Lessons learned and prospects. *Business Strategy and the Environment* 20(5): 331-347.
Bradach, J.L. 1997. Using the plural form in the management of restaurant chains. *Administrative Science Quarterly* 42(2): 276-303.
Bradach, J.L., and Eccles, R.G. 1989. Price, authority, and trust: From ideal types to plural forms. *Annual Review of Sociology* 15(1): 97-118.
Braithwaite, J., and Makkai, T. 1991. Testing an expected utility model of corporate deterrence. *Law & Society Review* 25(1): 7-40.
Brass, D.J., Butterfield, K.D., and Skaggs, B.C. 1998. Relationships and unethical behavior: A social network perspective. *Academy of Management Review* 23(1): 14-31.
Cameron, A.C. and Trivedi, P.K. 1998. *Regression Analysis of Count Data*, Cambridge University Press, New York.
Carcello, J.V., and Nagy, A.L. 2004. Audit firm tenure and fraudulent financial reporting. *Auditing: A Journal of Practice & Theory* 23(2): 55-69.
Causholli, M., Chambers, D.J., and Payne, J.L. 2014. Future nonaudit service fees and audit quality. *Contemporary Accounting Research* 31(3): 681-712.
Chugh, D., and Bazerman, M.H. 2007. Bounded awareness: What you fail to see can hurt you. *Mind and Society* 6(1): 1–18.
Coase, R.H. 1937. The nature of the firm. *Economica* 4(16): 386-405.
Coffee, J. C., Jr. 2004. Gatekeeper failure and reform: The challenge of fashioning relevant reforms. *Boston University Law Review* 84(2): 301-364.
Cressey, D.R., and Moore, C.A. 1983. Managerial values and corporate codes of ethics. *California Management Review* 25(4): 53-77.
Dao, M., Raghunandan, K., and Rama, D.V. 2012. Shareholder voting on auditor selection, audit fees, and audit quality. *Accounting Review* 87(1): 149-171.
Darnall, N., Ji, H., and Vázquez-Brust, D.A. 2018. Third-party certification, sponsorship, and consumers' ecolabel use. *Journal of Business Ethics* 150(4): 953-969.
DeAngelo, L.E. 1981. Auditor size and audit quality. *Journal of Accounting and Economics* 3(3): 183-199.
Duflo, E., Greenstone, M., Pande, R., and 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.
Dutta, S., Bergen, M., Heide, J.B, and John, G. 1995. Understanding dual distribution: The case of reps and house accounts. *Journal of Law, Economics, & Organization* 11(1): 189-204.
Dyck, I.J., Morse, A., and Zingales, L. 2017. How pervasive is corporate fraud? Rotman School of Management Working Paper 2222608.
Francis, J.R., and Wilson, E.R. 1988. Auditor changes: A joint test of theories relating to agency costs and auditor differentiation. *Accounting Review* 64(4): 663-682.
Gray, W.B., and Shadbegian, R.J. 2005. When and why do plants comply? Paper mills in the 1980s. *Law & Policy* 27(2): 238-261.
Greenberg, J., and Mollick, E. 2017. Activist choice homophily and the crowdfunding of female founders. *Administrative Science Quarterly* 62(2): 341-374.
Griffin, J.M. and Tang, D.Y. 2011. Did credit rating agencies make unbiased assumptions on CDOs? *American Economic Review 101*(3): 125-30.
Handley, S.M., and Benton, W.C., Jr. 2012. The influence of exchange hazards and power on

opportunism in outsourcing relationships. *Journal of Operations Management* 30(1-2): 55-68.
Harrigan, K.R. 1986. Matching vertical integration strategies to competitive conditions. *Strategic Management Journal 7*(6): 535-555.
Heide, J.B. 2003. Plural governance in industrial purchasing. *Journal of Marketing* 67(4): 18-29.
Heide, J.B., Kumar, A., and Wathne, K.H. 2014. Concurrent sourcing, governance mechanisms, and performance outcomes in industrial value chains. *Strategic Management Journal* 35(8): 1164-1185.
Iacus, S.M., King, G. and Porro, G. 2008. Matching for causal inference without balance checking. *Available at SSRN 1152391*.
Ibanez, M.R., and Toffel, M.W. 2020. How scheduling can bias quality assessment: Evidence from food-safety inspections. *Management Science* 66(6): 2396-2416.
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.
Jiang, J.X., Stanford, M.H., and Xie, Y. 2012. Does it matter who pays for bond ratings? Historical evidence. *Journal of Financial Economics* 105(3): 607-621.
Jin, G. Z., and Lee, J. (2018). A tale of repetition: Lessons from Florida restaurant inspections. *Journal of Law and Economics* 61(1): 159-188.
Khalil, F., and Lawarrée, J. 2006. Incentives for corruptible auditors in the absence of commitment. *Journal of Industrial Economics* 54(2): 269-291.
Kouakou, D., Boiral, O., and Gendron, Y. 2013. ISO auditing and the construction of trust in auditor independence. *Accounting, Auditing & Accountability Journal* 26(8): 1279-1305.
Kraakman, R.H. 1986. Gatekeepers: The anatomy of a third-party enforcement strategy. *Journal of Law, Economics, & Organization* 2(1): 53-104.
Krogh, G.V., Roos, J., and Slocum, K. 1994. An essay on corporate epistemology. *Strategic Management Journal* 15(S2): 53-71.
Kumar, J.A., and Chakrabarti, A. 2012. Bounded awareness and tacit knowledge: Revisiting Challenger disaster. *Journal of Knowledge Management* 16(6): 934-949.
Lebaron, G., and Lister, J. 2015. Benchmarking global supply chains: The power of the "ethical audit" regime. *Review of International Studies* 41(05): 905-924.
Lennox, C.S., Wu, X., and Zhang, T. 2014. Does mandatory rotation of audit partners improve audit quality? *The 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.
Ljungqvist, A., Marston, F., Starks, L.T., Wei, K.D., and 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.
Lundberg, S., and Stearns, J. 2019. Women in economics: Stalled progress. *Journal of Economic Perspectives* 33(1): 3-22.
Lytton, T.D., and McAllister, L.K. 2014. Oversight in private food safety auditing: Addressing auditor conflict of interest. *Wisconsin Law Review* 2014(2): 289-336.
Macher, J.T., and Richman, B.D. 2008. Transaction cost economics: An assessment of empirical research in the social sciences. *Business and Politics* 10(1): 1-63.
Markell, D.L., and Glicksman, R.L. 2014. A holistic look at agency enforcement. *North Carolina Law Review* 93(1): 1-18.
Mayer, K.J. 2006. Spillovers and governance: An analysis of knowledge and reputational spillovers in information technology. *Academy of Management Journal 49*(1): 69-84.
Mayhew, B.W., and Pike, J.E. 2004. Does investor selection of auditors enhance auditor

independence? *The Accounting Review* 79(3): 797-822.
McAllister, L.K. 2012. Regulation by third-party verification. *Boston College Law Review* 53(1): 1-32.
McBarnet, D. 2007. Corporate social responsibility beyond law, through law, for law: The new corporate accountability. In D. McBarnet, A. Voiculescu, and T. Campbell (eds.), *The New Corporate Accountability: Corporate Social Responsibility and the Law*: 9-56. Cambridge, UK: Cambridge University Press.
Metzger, M., Dalton, D.R., and Hill, J.W. 1993. The organization of ethics and the ethics of organizations: The case for expanded organizational ethics audits. *Business Ethics Quarterly* 3(1): 27-43.
Montiel, I., Husted, B.W., and Christmann, P. 2012. Using private management standard certification to reduce information asymmetries in corrupt environments. *Strategic Management Journal* 33(9): 1103–1113.
Moore, D.A., Tetlock, P.E., Tanlu, L., and Bazerman, M.H. 2006. Conflicts of interest and the case of auditor independence: Moral seduction and strategic issue cycling. *Academy of Management Review* 31(1): 10-29.
Nickerson, J.A., and Silverman, B.S. 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.
Oka, C. 2010. Accounting for the gaps in labour standard compliance: The role of reputation-conscious buyers in the Cambodian garment industry. *The European Journal of Development Research* 22(1): 59-78.
Parmigiani, A. 2007. Why do firms both make and buy? An investigation of concurrent sourcing. *Strategic Management Journal* 28(3): 285-311.
Parmigiani, A., and Mitchell, W. 2009. Complementarity, capabilities, and the boundaries of the firm: The impact of within-firm and interfirm expertise on concurrent sourcing of complementary components. *Strategic Management Journal* 30(10): 1065-1091.
Pierce L., and Toffel, M.W. 2013. The role of organizational scope and governance in strengthening private monitoring. *Organization Science* 24(5): 1558-1584.
Power, M. 1997. *The Audit Society: Rituals of Verification*. Oxford, UK: Oxford University Press.
Prakash, A., and Potoski, M. 2007. Collective action through voluntary environmental programs: A club theory perspective. *Policy Studies Journal* 35(4): 773-792.
Puranam, P., Gulati, R., and 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.
Rubineau, B., and Fernandez, R.M. 2015. Tipping points: The gender segregating and desegregating effects of network recruitment. *Organization Science* 26(6): 1646-1664.
Ruiz-Barbadillo, E., Gómez-Aguilar, N., and Carrera, N. 2009. Does mandatory audit firm rotation enhance auditor independence? Evidence from Spain. *Auditing: A Journal of Practice & Theory* 28(1): 113-135.
Sako, M., Chondrakis, G., and Vaaler, P.M. 2016. How do plural-sourcing firms make and buy? The impact of supplier portfolio design. *Organization Science* 27(5): 1161-1182.
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.
Short, J.L., Toffel, M.W., and Hugill, A.R. 2016. Monitoring global supply chains. *Strategic Management Journal* 37(9): 1878-1897.
Tepalagul, N., and Lin, L. 2015. Auditor independence and audit quality: A literature review. *Journal of Accounting, Auditing & Finance* 30(1): 101-121.

Treviño, L.K., Butterfield, K.D., and McCabe, D.L. 1998. The ethical context in organizations: Influences on employee attitudes and behaviors. *Business Ethics Quarterly* 8(3): 447-476.
Wang, K.J., and Tuttle, B.M. 2009. The impact of auditor rotation on auditor–client negotiation. *Accounting, Organizations and Society* 34(2): 222-243.
Wathne, K.H., and Heide, J.B. 2000. Opportunism in interfirm relationships: Forms, outcomes, and solutions. *Journal of Marketing* 64(4): 36-51.
Williamson, O. E. 1985. *The Economic Institutions of Capitalism*. New York: Macmillan.

Table 1: Summary statistics

	Mean	SD	Min	Max
New violations	6.26	6.69	0	76
Second-party auditor	0.36	0.48	0	1
Concurrent source	0.89	0.31	0	1
Auditor rotation				
Different audit firm	0.40	0.49	0	1
Different lead auditor (same audit firm)	0.19	0.39	0	1
Same lead auditor (same audit firm)	0.11	0.31	0	1
Missing rotation	0.29	0.46	0	1
Female lead auditor	0.40	0.46	0	1
Unannounced audit	0.17	0.37	0	1
Total workers	464	843	0	15,998
Total workers (log)	5.05	1.67	0	9.7
Audit type				
New factory audit	0.20	0.40	0	1
Limited scope audit	0.09	0.29	0	1
Re-audit	0.38	0.48	0	1
Follow-up audit	0.22	0.41	0	1
Factory tenure (years)	2.59	3.55	0	26
Percent supplied to MNC buyer	18.37%	26.99%	0	1
Audit sequence	3.15	2.16	1	8
Audit year	2014	2.1	2007	2017
N (audits)		15,935		

Table 2: Poisson regression results

Dependent variable = *new violations*		(1)		(2)		(3)	
		Coef	AME	Coef	AME	Coef	AME
H1	Second-party auditor	0.266**	1.9			0.197**	1.3
		(0.050)				(0.024)	
H2	Concurrent source			0.263**	1.2	0.041	0.3
				(0.088)		(0.055)	
H3	Different audit firm	0.298**	2.2	0.299**	1.6	0.236**	1.5
		(0.071)		(0.047)		(0.024)	
	Same lead auditor (same audit firm)	-0.391**	-2.3	-0.279**	-1.3	-0.401**	-2.1
		(0.108)		(0.056)		(0.038)	
	Female lead auditor	0.073	0.5	0.158**	0.8	0.150**	0.9
		(0.058)		(0.035)		(0.022)	
	Unannounced audit	0.348**	2.8	0.151**	0.8	0.120**	0.8
		(0.084)		(0.038)		(0.027)	
	Total workers (log)	0.080	0.6	0.091*	0.5	0.081**	0.5
		(0.070)		(0.045)		(0.026)	
	Limited scope audit	-0.925**	-4.5	-0.785**	-2.9	-0.674**	-3.2
		(0.244)		(0.090)		(0.076)	
	Re-audit	-0.258	-1.7	-0.069	-0.4	0.089	0.6
		(0.211)		(0.085)		(0.065)	
	Follow-up audit	-0.687**	-3.9	-0.518**	-2.4	-0.338**	-1.9
		(0.220)		(0.085)		(0.068)	
	Factory tenure	-0.043	-0.3	-0.033**	-0.2	-0.013	-0.1
		(0.051)		(0.013)		(0.020)	
	Percent supplied to MNC buyer	-0.067	-0.5	-0.007	-0.5	-0.085+	-0.5
		(0.159)		(0.041)		(0.051)	
	Limited scope audit (lag)	0.221	1.7	-0.093	1.7	-0.046	-0.3
		(0.219)		(0.060)		(0.063)	
	Re-audit (lag)	0.222	1.7	0.007	1.7	0.053	0.3
		(0.220)		(0.051)		(0.062)	
	Follow-up audit (lag)	0.249	1.9	-0.096	1.9	0.019	0.1
		(0.224)		(0.072)		(0.065)	
	Factory fixed effects	Yes		Yes		Yes	
	Audit-year fixed effects	Yes		Yes		Yes	
	Audit-sequence fixed effects	Yes		Yes		Yes	
	Sample	Matched sample		Third-party audits		All audits	
	N (focal audits)	3,526		10,071		15,935	
	Number of supplier factories	2,230		3,398		4,685	
	Sample average of *new violations*	7.0		5.2		6.3	

Notes: Poisson coefficients with robust standard errors clustered by supplier factory for Models 1 and 3 and by supplier factory's country for Model 2. + p<0.10, * p<0.05, **p<0.01. AME is average marginal effects. Omitted category for *different audit firm* and *same lead auditor (same audit firm)* is *different lead auditor (same audit firm).* Omitted category for audit type (*limited scope audit*, *re-audit*, and *follow-up audit*) is *new factory audit*.

APPENDIX

Table A-1: Correlations

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	New violations	1.00														
2	Second-party auditor	0.21	1.00													
3	Concurrent source	0.05	0.11	1.00												
4	Different audit firm	-0.12	-0.27	0.21	1.00											
5	Different lead auditor (same audit firm)	-0.17	-0.06	0.02	-0.40	1.00										
6	Same lead auditor (same audit firm)	-0.16	0.08	-0.19	-0.29	-0.17	1.00									
7	Missing rotation	-0.03	-0.03	0.00	-0.06	-0.04	-0.03	1.00								
8	Female lead auditor	0.00	-0.07	-0.12	-0.05	0.01	-0.01	0.00	1.00							
9	Unannounced audit	-0.04	0.05	0.08	0.16	0.04	0.07	0.00	-0.10	1.00						
10	Total workers (log)	0.08	0.03	0.13	0.03	0.01	0.02	0.01	0.09	-0.02	1.00					
11	New factory audit	0.34	0.25	-0.03	-0.40	-0.24	-0.17	-0.04	0.08	-0.19	0.02	1.00				
12	Limited scope audit	-0.12	0.22	0.02	-0.05	0.13	0.20	0.00	0.02	0.00	0.08	-0.16	1.00			
13	Re-audit	-0.12	-0.33	0.01	0.30	0.12	0.02	-0.01	0.03	0.09	-0.07	-0.38	-0.25	1.00		
14	Follow-up audit	-0.17	-0.08	0.12	0.21	0.09	0.01	0.06	-0.07	0.19	-0.02	-0.26	-0.17	-0.41	1.00	
15	Factory tenure (years)	-0.16	-0.18	0.00	0.19	0.10	0.04	0.01	0.09	-0.34	0.06	-0.19	0.22	0.20	-0.06	1.00
16	Percent supplied to MNC buyer	-0.12	-0.16	0.08	0.22	0.04	-0.01	0.03	0.12	-0.28	0.02	-0.17	0.19	0.19	-0.04	0.29

Table A-2. Summary statistics for samples used in primary tests of H1–H2 (Models 1 and 2 of Table 2)

	Model 1 sample to test H1				Model 2 sample to test H2			
	Mean	SD	Min	Max	Mean	SD	Min	Max
New violations	7.00	6.88	0	67	5.19	5.71	0	76
Second-party auditor	0.50	0.50	0	1	0.00	0.00		
Concurrent source	0.99	0.10	0	1	0.87	0.34	0	1
Auditor rotation								
Different audit firm	0.37	0.48	0	1	0.50	0.50	0	1
Different lead auditor (same audit firm)	0.14	0.34	0	1	0.21	0.41	0	1
Same lead auditor (same audit firm)	0.07	0.26	0	1	0.09	0.29	0	1
Missing rotation	0.41	0.49	0	1	0.20	0.40	0	1
Female lead auditor	0.43	0.47	0	1	0.43	0.46	0	1
Unannounced audit	0.15	0.36	0	1	0.15	0.36	0	1
Total workers	485	954	0	15,998	407	769	0	15,998
Total workers (log)	5.10	1.59	0	9.7	5.00	1.57	0	9.7
Audit type								
New factory audit	0.30	0.46	0	1	0.12	0.33	0	1
Limited scope audit	0.10	0.30	0	1	0.04	0.20	0	1
Re-audit	0.28	0.45	0	1	0.50	0.50	0	1
Follow-up audit	0.18	0.38	0	1	0.24	0.43	0	1
Factory tenure	2.36	3.56	0	26	3.06	3.62	0	25
Percent supplied to MNC buyer	0.17	0.26	0	1	0.22	0.28	0	1
Audit sequence	3.01	2.40	1	8	3.49	2.11	1	8
N (audits)	3,526				10,071			

Table A-3: Covariate balance of matched sample to test H1 in Table 2 Model 1

	Total sample				Matched sample			
	Second-party mean	Third-party mean	p-value	Std bias	Second-party mean	Third-party mean	p-value	Std bias
New violations (prior audit)	4.35	4.18	0.09	2.67	3.17	3.20	0.85	-0.62
Total violations (prior audit)	6.07	6.62	0.00	-7.39	5.14	5.19	0.80	-0.86
Social audit rating (prior audit)								
Red-Critical rating (prior audit)	0.09	0.03	0.00	26.47	0.02	0.02	1.00	0.00
Red rating (prior audit)	0.28	0.28	0.94	-0.13	0.28	0.28	1.00	0.00
Amber rating (prior audit)	0.14	0.43	0.00	-66.85	0.27	0.27	1.00	0.00
Green rating (prior audit)	0.01	0.05	0.00	-24.65	0.01	0.01	1.00	0.00
Missing rating (prior audit)	0.48	0.21	0.00	58.56	0.42	0.42	1.00	0.00
Concurrent source	0.93	0.86	0.00	23.27	0.99	0.99	1.00	0.00
Different audit firm	0.23	0.50	0.00	-58.71	0.31	0.44	0.00	-27.19
Different lead auditor (same audit firm)	0.16	0.21	0.00	-12.26	0.17	0.10	0.00	20.50
Same lead auditor (same audit firm)	0.14	0.09	0.00	16.16	0.11	0.04	0.00	24.48
Second-party auditor (prior audit)	0.31	0.21	0.00	20.91	0.28	0.25	0.04	6.98
Female lead auditor	0.36	0.43	0.00	-15.55	0.44	0.43	0.54	2.07
Announcement								
Unannounced audit	0.19	0.15	0.00	10.44	0.15	0.15	1.00	0.00
Missing announcement	0.09	0.03	0.00	25.43	0.05	0.05	1.00	0.00
Total workers (log)	5.12	5.01	0.00	6.93	5.08	5.12	0.47	-2.44
Total workers	527	394	0.00	15.72	458	480	0.50	-2.25
Audit scope								
New factory audit	0.33	0.12	0.00	51.29	0.30	0.30	1.00	0.00
Limited scope audit	0.18	0.04	0.00	43.26	0.10	0.10	1.00	0.00
Re-audit	0.16	0.50	0.00	-76.63	0.28	0.28	1.00	0.00
Follow-up audit	0.17	0.24	0.00	-16.59	0.18	0.18	1.00	0.00
Missing audit scope	0.16	0.10	0.00	19.27	0.13	0.13	1.00	0.00
Factory tenure	1.76	3.07	0.00	-38.10	2.32	2.40	0.47	-2.41
Percentage supplied to MNC buyer	0.13	0.22	0.00	-33.48	0.17	0.17	0.57	1.89
Audit scope (prior audit)								
New factory audit (prior audit)	0.15	0.13	0.00	5.92	0.11	0.11	0.79	0.91
Limited scope audit (prior audit)	0.05	0.06	0.00	-5.89	0.04	0.05	0.47	-2.46
Follow-up audit (prior audit)	0.15	0.33	0.00	-42.29	0.23	0.23	0.78	0.95
Re-audit (prior audit)	0.11	0.18	0.00	-21.89	0.13	0.14	0.62	-1.65
Missing audit scope (prior audit)	0.55	0.30	0.00	51.58	0.48	0.48	0.81	0.79
Audit sequence	2.56	3.48	0.00	-43.36	3.01	3.01	1.00	0.00
Audit year	2014	2014	0.27	-1.76	2014	2014	1.00	0.00
N (focal audits)		15,935				3,526		
No. of unbalanced variables		30				4		
Mean standardized bias		-3.48				0.58		
Median standardized bias		-0.95				0.00		

Notes: p-values obtained from t-tests for continuous variables and proportion tests for binary variables. Standardized bias is calculated by dividing the mean difference in group means by the average standard deviation. The matched sample was found by exactly matching second-party audits to third-party audits on supplier country, *audit sequence*, *audit year, social audit rating (prior audit), concurrent source, audit scope,* and *audit announcement* and by coarsened exact matching on *new violations (prior audit).*

Table A-4. Poisson regression results of robustness tests for H1

	(1)	(2)	(3)	(4)
Dependent variable:	New violations	New violations top-coded at 95th percentile	Major new violations	Total violations
Sample:	Full data sample	Matched sample	Matched sample	Matched sample
	Coef	Coef	Coef	Coef
Second-party auditor	0.151**	0.248**	0.264**	0.151**
	(0.032)	(0.045)	(0.053)	(0.032)
Different audit firm	0.109**	0.308**	0.307**	0.109**
	(0.042)	(0.067)	(0.076)	(0.042)
Same lead auditor (same audit firm)	-0.202**	-0.378**	-0.354**	-0.202**
	(0.057)	(0.096)	(0.119)	(0.057)
Female lead auditor	0.012	0.062	0.057	0.012
	(0.033)	(0.056)	(0.060)	(0.033)
Unannounced audit	0.140*	0.314**	0.377**	0.140*
	(0.055)	(0.078)	(0.088)	(0.055)
Total workers (log)	0.020	0.061	0.054	0.020
	(0.041)	(0.066)	(0.075)	(0.041)
Limited scope audit	-0.449**	-0.865**	-1.047**	-0.449**
	(0.172)	(0.217)	(0.260)	(0.172)
Re-audit	-0.147	-0.243	-0.344	-0.147
	(0.167)	(0.191)	(0.227)	(0.167)
Follow-up audit	-0.286+	-0.651**	-0.803**	-0.286+
	(0.170)	(0.198)	(0.238)	(0.170)
Factory tenure	-0.076+	-0.039	-0.045	-0.076+
	(0.044)	(0.047)	(0.054)	(0.044)
Percent supplied to MNC buyer	-0.054	-0.035	-0.097	-0.054
	(0.100)	(0.142)	(0.173)	(0.100)
Limited scope audit (lag)	-0.031	0.255	0.132	-0.031
	(0.150)	(0.202)	(0.245)	(0.150)
Re-audit (lag)	0.032	0.279	0.109	0.032
	(0.146)	(0.203)	(0.246)	(0.146)
Follow-up audit (lag)	0.009	0.325	0.120	0.009
	(0.155)	(0.202)	(0.252)	(0.155)
Factory fixed effects	Yes	Yes	Yes	Yes
Audit-year fixed effects	Yes	Yes	Yes	Yes
Audit-sequence fixed effects	Yes	Yes	Yes	Yes
N (audits)	15,935	3,526	3,526	3,526
Number of supplier factories	4,685	2,230	2,230	2,230
Sample average of dependent variable	6.3	6.5	6.0	9.8

Notes: Poisson coefficients clustered by supplier factory. + p<.10, * p<.05, **p<.01. Omitted category for *different audit firm* and *same lead auditor (same audit firm)* is *different lead auditor (same audit firm)*. Omitted category for *limited scope audit*, *re-audit*, and *follow-up audit* is *new factory audit*.

Table A-5. Poisson regression results based on alternative concurrent sourcing thresholds to test H2

Dependent variable: *new violations*	(1)	(2)	(3)	(4)	(5)	(6)
Thresholds to define concurrent sourcing:	1-99%	5-95%	10-90% (primary approach)	15-85%	20-80%	25-75%
Concurrent source	0.336**	0.299**	0.263**	0.239*	0.237**	0.230**
	(0.108)	(0.094)	(0.088)	(0.106)	(0.068)	(0.070)
Different audit firm	0.297**	0.305**	0.299**	0.291**	0.293**	0.269**
	(0.049)	(0.051)	(0.047)	(0.044)	(0.047)	(0.051)
Same lead auditor (same audit firm)	-0.266**	-0.258**	-0.279**	-0.294**	-0.283**	-0.307**
	(0.051)	(0.055)	(0.056)	(0.058)	(0.055)	(0.054)
Female lead auditor	0.153**	0.159**	0.158**	0.153**	0.152**	0.177**
	(0.037)	(0.035)	(0.035)	(0.037)	(0.043)	(0.044)
Unannounced audit	0.151**	0.154**	0.151**	0.151**	0.149**	0.156**
	(0.039)	(0.037)	(0.038)	(0.039)	(0.036)	(0.034)
Total workers (log)	0.079	0.080	0.091*	0.094*	0.083*	0.107*
	(0.048)	(0.049)	(0.045)	(0.045)	(0.041)	(0.050)
Limited scope audit	-0.800**	-0.788**	-0.785**	-0.813**	-0.838**	-0.776**
	(0.089)	(0.086)	(0.090)	(0.086)	(0.094)	(0.141)
Re-audit	-0.051	-0.044	-0.069	-0.100	-0.133+	-0.049
	(0.095)	(0.094)	(0.085)	(0.075)	(0.077)	(0.145)
Follow-up audit	-0.495**	-0.484**	-0.518**	-0.545**	-0.590**	-0.499**
	(0.098)	(0.097)	(0.085)	(0.084)	(0.089)	(0.146)
Factory tenure	-0.027*	-0.028**	-0.033**	-0.023*	-0.027*	-0.022
	(0.011)	(0.011)	(0.013)	(0.010)	(0.011)	(0.018)
Percent supplied to MNC buyer	-0.004	0.005	-0.007	-0.016	-0.065	0.018
	(0.046)	(0.043)	(0.041)	(0.045)	(0.051)	(0.054)
Limited scope audit (prior audit)	-0.097	-0.103	-0.093	-0.102+	-0.069	-0.056
	(0.062)	(0.063)	(0.060)	(0.058)	(0.069)	(0.071)
Re-audit (prior audit)	0.005	0.007	0.007	-0.001	0.016	0.003
	(0.057)	(0.057)	(0.051)	(0.049)	(0.059)	(0.091)
Follow-up audit (prior audit)	-0.093	-0.089	-0.096	-0.104	-0.111	-0.106
	(0.073)	(0.075)	(0.072)	(0.074)	(0.078)	(0.110)
Factory fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Audit-year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Audit-sequence fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
N (audits)	10,061	10,056	10,071	10,060	9,295	9,080
Supplier factories	3,403	3,402	3,398	3,396	3,344	3,306
Sample average of dependent variable	5.2	5.2	5.2	5.2	5.2	5.2

Notes: Poisson coefficients with robust standard errors clustered at the supplier factory country. + p<0.10, * p<0.05, **p<0.01. To isolate the effect of transitioning from a third-party market to a concurrent sourcing market, we exclude from each sample the few instances in which the reverse transition occurred. Omitted category for *different audit firm* and *same lead auditor (same audit firm)* is *different lead auditor (same audit firm)*. Omitted category for *limited scope audit*, *re-audit*, and *follow-up audit* is *new factory audits*.

Table A-6. Poisson regression results of robustness tests for H2

Dependent variable:	(1) New violations top-coded at 95th percentile	(2) Major new violations	(3) Total violations
Concurrent source	0.283**	0.236*	0.140*
	(0.067)	(0.104)	(0.055)
Different audit firm	0.287**	0.321**	0.146**
	(0.035)	(0.050)	(0.035)
Same lead auditor (same audit firm)	-0.265**	-0.288**	-0.085+
	(0.056)	(0.066)	(0.045)
Female lead auditor	0.144**	0.158**	0.073+
	(0.030)	(0.047)	(0.042)
Unannounced audit	0.139**	0.154**	0.097**
	(0.042)	(0.039)	(0.031)
Total workers (log)	0.054+	0.075+	0.000
	(0.032)	(0.041)	(0.029)
Limited scope audit	-0.742**	-0.816**	-0.456**
	(0.081)	(0.087)	(0.085)
Re-audit	-0.087	-0.057	-0.139+
	(0.070)	(0.093)	(0.077)
Follow-up audit	-0.520**	-0.524**	-0.262**
	(0.075)	(0.087)	(0.086)
Factory tenure	-0.050**	-0.028*	-0.031
	(0.011)	(0.011)	(0.019)
Percent supplied to MNC buyer	-0.002	-0.020	-0.036
	(0.043)	(0.044)	(0.026)
Limited scope audit (prior audit)	-0.099**	-0.144*	-0.203**
	(0.038)	(0.066)	(0.059)
Re-audit (prior audit)	0.009	-0.018	-0.053
	(0.042)	(0.059)	(0.048)
Follow-up audit (prior audit)	-0.069	-0.137	-0.133*
	(0.061)	(0.097)	(0.063)
Factory fixed effects	Yes	Yes	Yes
Audit-year fixed effects	Yes	Yes	Yes
Audit-sequence fixed effects	Yes	Yes	Yes
N (audits)	10,071	10,071	10,071
Supplier factories	3,398	3,398	3,398
Sample average of dependent variable	4.9	4.4	8.7

Notes: Poisson coefficients with robust standard errors clustered at the supplier factory country. + p<0.10, * p<0.05, **p<0.01. To isolate the effect of transitioning from a third-party market to a concurrent sourcing market, we exclude from each sample the few instances in which the reverse transition occurred. Omitted category for *different audit firm* and *same lead auditor (same audit firm)* is *different lead auditor (same audit firm)*. Omitted category for *limited scope audit*, *re-audit*, and *follow-up audit* is *new factory audits*.

Table A-7. Poisson regression results of robustness tests for H3

	(1)	(2)	(3)	(4)	(5)
Dependent variable:	New violations	New violations	New violations top-coded at 95th percentile	Major new violations	Total violations
Sample:	See notes	Matched sample	All audits	All audits	All audits
Second-party auditor	0.197**	0.234**	0.185**	0.201**	0.125**
	(0.026)	(0.047)	(0.021)	(0.026)	(0.016)
Concurrent source	0.012	0.154	0.067	-0.046	-0.049
	(0.059)	(0.116)	(0.044)	(0.063)	(0.041)
Different audit firm	0.223**	0.297**	0.229**	0.254**	0.108**
	(0.027)	(0.034)	(0.022)	(0.026)	(0.015)
Same lead auditor (same audit firm)	-0.433**	-0.284**	-0.370**	-0.407**	-0.156**
	(0.043)	(0.064)	(0.034)	(0.043)	(0.024)
Female lead auditor	0.150**	0.183**	0.126**	0.132**	0.056**
	(0.024)	(0.035)	(0.020)	(0.023)	(0.013)
Unannounced audit	0.134**	0.173**	0.110**	0.136**	0.088**
	(0.030)	(0.047)	(0.023)	(0.028)	(0.017)
Total workers (log)	0.100**	0.115*	0.065**	0.074**	-0.000
	(0.032)	(0.047)	(0.018)	(0.026)	(0.015)
Limited scope audit	-0.811**	-0.764**	-0.592**	-0.693**	-0.415**
	(0.083)	(0.091)	(0.065)	(0.082)	(0.056)
Re-audit	-0.015		0.077	0.098	-0.043
	(0.072)		(0.055)	(0.070)	(0.052)
Follow-up audit	-0.455**	-0.426**	-0.331**	-0.340**	-0.176**
	(0.074)	(0.035)	(0.059)	(0.075)	(0.054)
Factory tenure	-0.003	-0.075*	-0.031+	-0.025	-0.015
	(0.023)	(0.035)	(0.016)	(0.021)	(0.019)
Percent supplied to MNC buyer	-0.068	0.031	-0.037	-0.100+	-0.109**
	(0.057)	(0.098)	(0.045)	(0.056)	(0.040)
Limited scope audit (prior audit)	-0.131+	-0.111	-0.017	-0.084	-0.156**
	(0.077)	(0.071)	(0.057)	(0.068)	(0.043)
Re-audit (prior audit)	-0.052		0.069	0.044	0.001
	(0.068)		(0.056)	(0.066)	(0.040)
Follow-up audit (prior audit)	-0.081	-0.115**	0.054	-0.000	-0.060
	(0.072)	(0.036)	(0.058)	(0.070)	(0.043)
Factory fixed effects	Yes	Yes	Yes	Yes	Yes
Audit-year fixed effects	Yes	Yes	Yes	Yes	Yes
Audit-sequence fixed effects	Yes	Yes	Yes	Yes	Yes
N (audits)	13,987	6,046	15,935	15,935	15,935
Number of supplier factories	4,685	2,231	4,685	4,685	4,685
Sample average of dependent variable	6.4	4.4	5.9	5.3	9.6

Notes: The sample of the model reported in Column 1 excludes audits in which the factory's improvement score, calculated as the percent reduction in number of last-audit *new violations* from the average *new violations* recorded in a factory's preceding two audits, is within the largest 25th percentile of improvement scores.

Poisson coefficients with robust standard errors clustered at the supplier. + p<0.10, * p<0.05, **p<0.01. Omitted category for *different audit firm* and *same lead auditor (same audit firm)* is *different lead auditor (same audit firm)*. Omitted category for audit type (*limited scope audit*, *re-audit*, and *follow-up audit)* is *new factory audit,* except for Column 2, where *re-audit* is the omitted category.

Table A-8. Poisson regression results to assess the impact of familiarity on audit firm rotation

Dependent variable = *new violations*	(1)	
	Coef	AME
New different audit firm	0.265**	1.8
	(0.026)	
Familiar different audit firm	0.155**	1.0
	(0.031)	
Same lead auditor (same audit firm)	-0.398**	-2.1
	(0.038)	
Second-party auditor	0.209**	1.3
	(0.025)	
Concurrent source	0.034	0.2
	(0.055)	
Female lead auditor	0.149**	0.9
	(0.022)	
Unannounced audit	0.120**	0.8
	(0.027)	
Total workers (log)	0.082**	0.5
	(0.026)	
Limited scope audit	-0.662**	-3.2
	(0.077)	
Re-audit	0.097	0.6
	(0.065)	
Follow-up audit	-0.327**	-1.8
	(0.069)	
Factory tenure	-0.013	-0.1
	(0.020)	
Percent supplied to MNC buyer	-0.084	-0.5
	(0.051)	
Limited scope audit (prior audit)	-0.050	-0.3
	(0.063)	
Re-audit (prior audit)	0.056	0.2
	(0.062)	
Follow-up audit (prior audit)	0.029	0.4
	(0.065)	
Factory fixed effects	Yes	
Audit-year fixed effects	Yes	
Audit-sequence fixed effects	Yes	
N (audits)	15,935	
Supplier factories	4,685	
Sample average of dependent variable	6.3	

Notes: Poisson coefficients with robust standard errors clustered at the supplier. + p<0.10, * p<0.05, **p<0.01. Omitted category for *different audit firm* and *same lead auditor (same audit firm)* is *different lead auditor (same audit firm)*. Omitted category for *limited scope audit*, *re-audit*, and *follow-up audit* is *new factory audits*. AME is average marginal effects.

Figure A-1: Kernel density graphs illustrating balance in the primary matched sample used to test H1

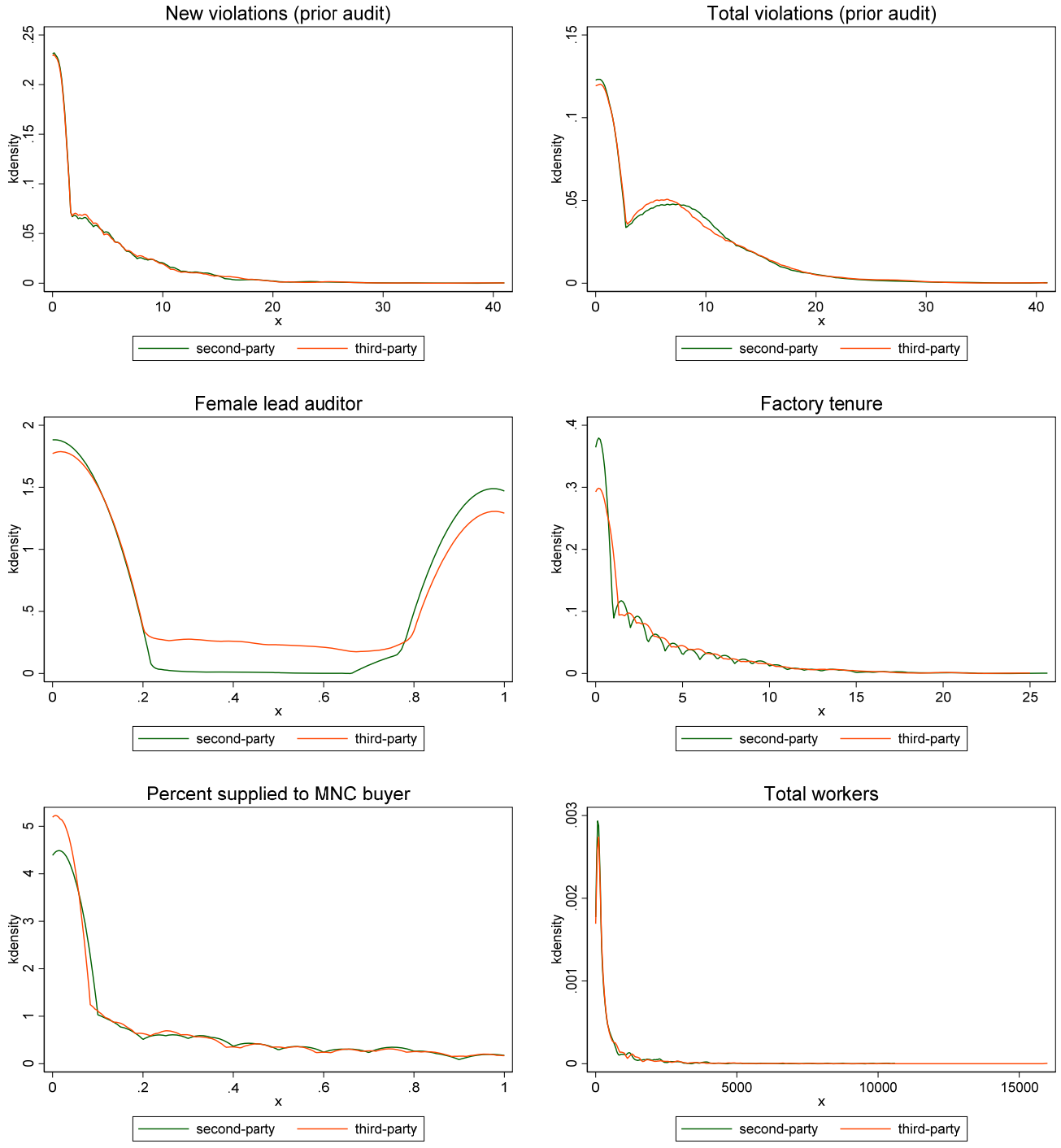


Figure A-1 (continued)

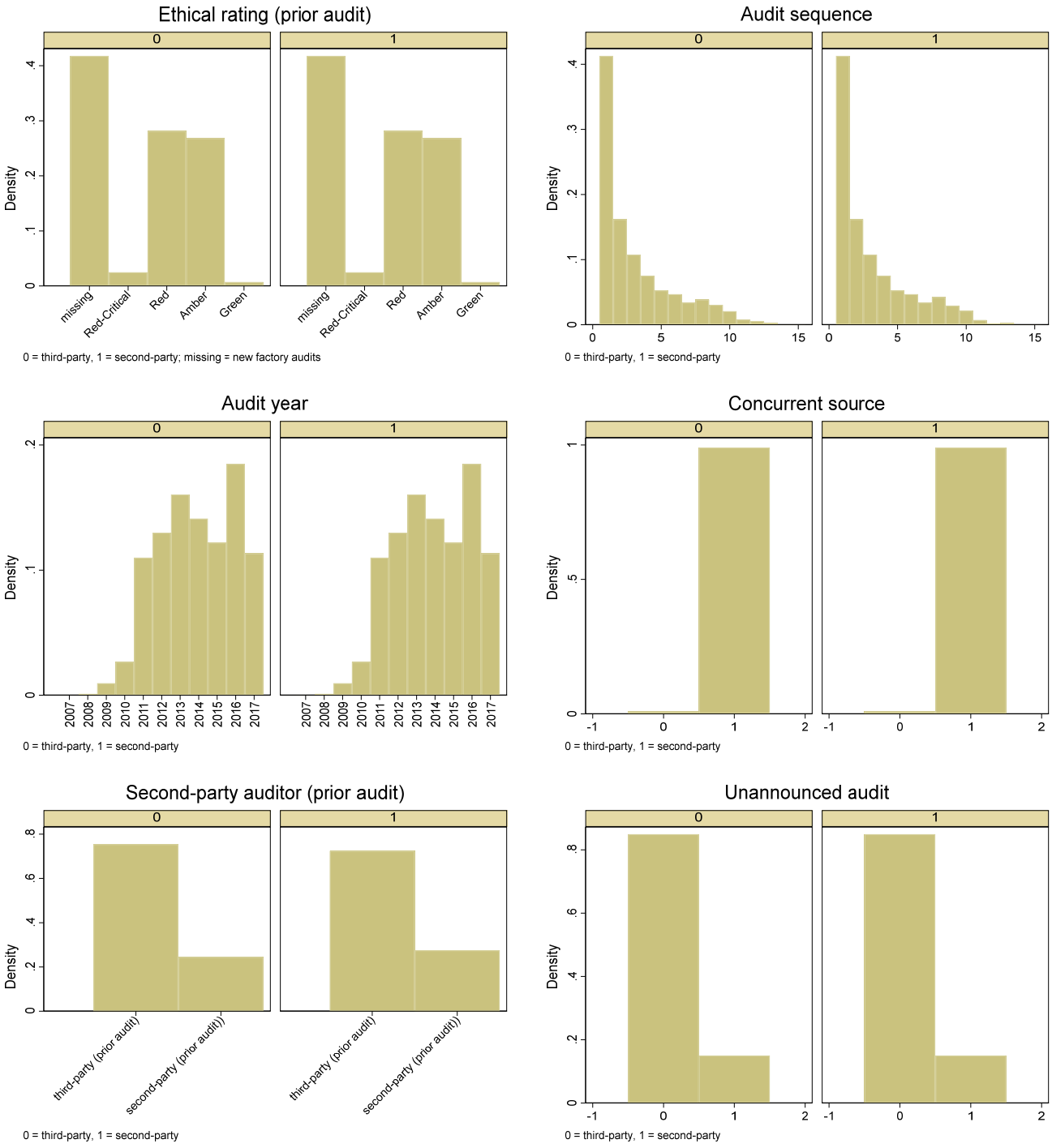


Figure A-1 (continued)