Working Paper 24-002

Managing Remote Work Quality: Evidence from Auditing Management Systems Standards

Ashley Palmarozzo
Michael W. Toffel

Harvard Business School

Managing Remote Work Quality: Evidence from Auditing Management Systems Standards

Ashley Palmarozzo
Harvard Business School

Michael W. Toffel
Harvard Business School

Working Paper 24-002

We gratefully acknowledge the research assistance provided by Melissa Ouellet.

Funding for this research was provided in part by Harvard Business School.

Managing Remote Work Quality:
Evidence from Auditing Management Systems Standards

Ashley Palmarozzo
Harvard Business School
Boston MA 02163
ashley.palmarozzo@gmail.com

Michael W. Toffel[†]
Harvard Business School
Boston MA 02163
mtoffel@hbs.edu

August 13, 2024

Remote work has become more common, providing operational flexibility and productivity benefits, but questions remain about whether and how it affects work quality. We investigate the quality effects of remote work in a context in which remote work separates workers from the subject of their work: remotely auditing compliance with management systems standards. We analyze nearly 35,000 audits conducted in-person or remotely across thousands of sites around the world by auditors of one global auditing company during 2019–2021, when remote auditing accelerated due to COVID-19 pandemic travel restrictions. We theorize that remote audits will be of lower quality (less comprehensive) than in-person audits because remote auditors face greater difficulties (a) obtaining information critical to detecting violations and (b) coordinating actions and exchanging information with fellow auditors. We find evidence of these theorized mechanisms: remote audits report fewer violations and quality problems are especially pronounced for (a) standards clauses in which auditors assess compliance via direct observation as opposed to document review and (b) audits conducted by multi-auditor teams. We find the quality problems of remote audits partially mitigated when the auditors had previously conducted in-person audits of the site or when the auditors' experience was more concentrated on the standard being audited. Understanding mechanisms by which remote work can erode quality and revealing attributes that exacerbate or attenuate such quality concerns can help companies and regulatory agencies better manage remote work.

Keywords: remote work, quality, audits, inspections, standards

1. Introduction

Remote work has become increasingly common across industries as varied as healthcare, education, information technology, and financial services (US Bureau of Labor Statistics 2022). Many expect remote work, accelerated by the COVID-19 pandemic, to endure and perhaps constitute 20% of workdays in the post-pandemic US economy (Barrero et al. 2021). Remote work has been shown to improve job satisfaction and productivity and reduce costs, but uncertainty remains about how it affects work quality. The few studies of the quality implications of remote work have compared situations—responding to

---

[†] Corresponding author. We gratefully acknowledge the research assistance provided by Melissa Ouellet and the financial support from Harvard Business School's Division of Research and Faculty Development.

phone calls or reviewing documents—in which workers changed their locations but not their work processes (Bloom et al. 2015, Choudhury et al. 2021).

But remote work often does change how work is conducted, with in-person interactions becoming technology-mediated via telephone, video, and online chats, as with bank tellers, customer service desks, healthcare appointments, and regulatory inspections. Moreover, when workers can work from home or "from anywhere," there is physical separation not only between service providers and their customers, but also between the service providers themselves. Both factors can reduce work quality.

We examine the quality implications of remote work in the context of auditing management systems standards, a service context that—much like social audits of global supply chain factories, quality audits of franchisees to assess contract compliance, and regulatory inspections assessing occupational health and safety, environmental, food safety, and process quality practices—has traditionally involved onsite visits during which one or more auditors conduct site tours to observe operations, interview employees face-to-face, and review physical documents. When conducting audits remotely, these interactions are mediated by technology such as telephones, video conferencing, and document-sharing applications. We study how the shift from in-person to remote work affects the quality of *management systems standard certification assessments*, which are audits of the extent to which a site has implemented procedures required by management systems standards.

Our research is timely, given the ongoing debate about the quality of remote workers conducting audits and regulatory inspections. Enthusiasts include both certification auditors and audited businesses. 80% of them indicated in a 2021 survey that "remote procedures give the same confidence as onsite audits" and the majority were satisfied with their remote audit experience and wanted to see "continued or increased use of remote activities" (IAF, ILAC, and ISO [International Accreditation Forum, International Laboratory Accreditation Cooperation, and International Organization for Standardization] 2021: 1–2). But this is not a consensus. A survey of environment, health, and safety (EHS) professionals indicated that while most had shifted at least some EHS audits from in-person to remote, "concern over the quality of results was by far the greatest impediment to starting a virtual auditing program"—in particular, concern that remote audits might yield "less thorough/robust results" (NAEM 2021: 18). Another survey found internal auditors concerned that auditing their own company's operations remotely makes it more difficult to access information and produce quality audits (Internal Audit Foundation and AuditBoard 2021). Governments, too, are concerned; a US Department of Labor (2021: 2) report suggests that remote inspections conducted by the Occupational Safety and Health Administration during the COVID-19 pandemic risked hazards going "unidentified and unabated longer."

We theorize that two mechanisms will lead remote auditing to produce lower-quality (less-comprehensive) audits by limiting an auditor's ability to gather sufficient evidence of violations, leading them to document fewer violations than are truly occurring. First, auditors face greater difficulty gathering evidence physically embedded in the site, such as slippery floors. Remote auditing replaces in-person site tours with video-mediated site tours, which do not convey certain contextual information such as smells, sounds, and fine-grained visual clues. Suspicious smells cannot be detected over Zoom, for example. Second, it becomes harder for auditors to access information gathered by fellow audit team members. Onsite audit team members frequently exchange information throughout the audit, coordinate their activities, and assess their progress in real time. We theorize that there will be less of this if auditors aren't actually onsite together, making it harder for them to piece together clues of noncompliance.

We also theorize that information-access challenges associated with remote auditing will be attenuated—and thus the quality decrement of remote audits reduced—when auditors have greater prior in-person exposure to the audited site. Such auditors will be able to rely on experience-based heuristics that can prompt important questions even in the absence of onsite cues.

To test our theory, we obtained data from a large company that audits sites around the world for compliance with management systems standards. Specifically, we examine the tens of thousands of audits conducted in-person or remotely during 2019–2021 for six of the world's most popular management systems standards: ISO 9001 Quality Management Systems standard, ISO 14001 Environmental Management Systems standard, ISO 27001 Information Technology Management Systems standard, OHSAS 18001 and ISO 45001 Occupational Health and Safety Management Systems standards, and ISO

13485 Medical Devices Quality Management Systems standard. Sixty-one percent of these audits were conducted fully in-person and 39% fully remotely. Most of the audited sites were in China, India, the United Kingdom, and the United States; the rest were in nearly 100 other countries.

Our empirical analysis indicates that remote audits yielded an average of 25% (or 0.40) fewer violations than in-person audits and we found evidence of our two theorized mechanisms. First, we investigated whether remote auditing quality suffers due to greater difficulty accessing the information that leads to citations—information primarily gathered through direct observation during site tours. Exploiting the fact that auditors rely on different auditing methods as the primary mechanism to identify violations of particular management systems standard clauses, we examine whether the quality deficit of remote audits is especially pronounced for those clauses for which auditors tend to identify violations via direct observation—that is, where remote auditors are at a particular disadvantage compared to in-person auditors—versus those clauses for which auditors tend to identify violations via document review—where remote and in-person access is more similar. Consistent with our theory, we find that the audit-quality decrement for remote audits is significantly larger for clauses for which auditors typically rely on direct observation. Second, we find that remote auditing quality concerns are exacerbated when audits are conducted by teams rather than by a single auditor. This is consistent with the second theorized mechanism: remote auditing makes it harder for auditors to coordinate and to share information so as to comprehensively identify violations.

We find two factors—both auditor attributes—that partially attenuate quality concerns of remote auditing. Specifically, we find that such concerns are less severe for audits conducted by auditors with (a) more prior in-person auditing exposure to the audited site or (b) more concentrated auditing experience with the management systems standard being audited.

Our work contributes to the remote work literature by finding evidence that working remotely can lower quality. While prior research found no quality difference in remote work conducted by telephone customer service representatives or patent examiners reviewing documents (Bloom et al. 2015, Choudhury et al. 2021), remote workers in those contexts are no more separated from the callers or the patent documents than when they work in the office. Our context, however, separates auditors from in-person engagement with audited sites and with audit team members. Our results, taken together, suggest that the quality implications of remote work depend on the extent to which working remotely changes work processes. Managers should therefore be especially careful to guard against quality loss when remotely conducting work that traditionally has required gathering information in-person, such as medical exams and safety inspections.

Our research also contributes to the literature on the quality of monitoring. While there is significant scholarship on audit-, auditor-, and institution-level factors that affect audit quality for in-person audits, our study is, to our knowledge, the first to examine whether varying the audit format affects quality. It also offers insights for monitoring quality in other domains, such as assessing retail franchisees' compliance with franchiser terms and global supply chain factories' adherence to brands' supplier codes of conduct. Understanding the relationship between audit format and audit quality is especially timely given the growing interest in remote auditing in regulatory, supply chain, and standard certification contexts (IAF, ILAC, and ISO 2021, Mofid et al. 2021, US Department of Health and Human Services 2024). Given that remote auditing does reduce costs and travel time and increases flexibility, our results can guide audit providers on how to staff audits to minimize the associated quality concerns; for example, by focusing remote audits on less-risky sites or on those that require only one auditor. Moreover, as explained in the Discussion section, audit providers can use our results to develop a hybrid approach; for example, remote document review combined with onsite observation.

2. Related Literature

Our work relates to the literature comparing in-person and remote work and to the literature on the effectiveness of monitoring the management practices of companies' business partners.

2.1. Remote Work

Remote work has been found to lead to higher worker productivity and job satisfaction and lower attrition (Bloom et al. 2015, Choudhury et al. 2021)[1] and to more-efficient labor allocation and improved customer service (Sun et al. 2020). Banks, retailers, and hospitals offering online options to supplement in-person service has been found to lead to increased customer retention, greater customer use of *offline* service channels, and more overall customer interactions (Buell et al. 2010, Campbell and Frei 2010, Xue et al. 2011, Bavafa et al. 2018, Bell et al. 2018, Rajan et al. 2019).[2]

Closer to our research are studies of how working remotely affects work quality; these have yielded mixed results. No difference in quality was found between call center customer service representatives who worked from home versus at the office (Bloom et al. 2015) or between patent examiners working from home versus "working from anywhere" (Choudhury et al. 2021).[3] But Bettinger et al. (2017) found worse educational outcomes for students who took a course online rather than in person, theorizing that the online students had less instructor oversight and were thus less motivated. Sun et al. (2020) found that hospital emergency room physicians remotely delivering healthcare services improved one dimension of service quality—reducing patient length-of-stay, both by reducing waiting time (treatment delays) and by more flexibly allocating the (remote) physicians to patients—with no erosion in care quality (readmission rates and in-hospital mortality rates). Another healthcare study found better patient health outcomes when doctors supplemented in-person visits with virtual channels to communicate with patients (Bavafa et al. 2018).

Our study also examines the quality implications of remote work, but differs along two dimensions. First, prior studies theorize that differences in the quality of remote versus in-person work

---

[1] Tan and Netessine (2020) also find that introducing tabletop technology in restaurants to assist in ordering increased waiter productivity and sales per table.

[2] One study examining the reverse transition found that when a retailer increased sales conversion and decreased product returns when it supplemented its online store with in-person showrooms (Bell et al. 2018).

[3] Specifically, Bloom et al. (2015) found no difference in sales conversion rates or quality assessments for the recorded conversations and Choudhury et al. (2021) found no difference in rework or the number of citations patent examiners added during their patent reviews.

result from differences in service delays (Sun et al. 2020) and oversight (Bettinger et al. 2017), whereas these dimensions do not meaningfully differ between remote versus in-person work in our auditing context. Second, prior studies that compare the quality of office work performed remotely versus in-person consider scenarios in which worker location differed but the work itself did not entail in-person interaction with customers or colleagues either way. When Bloom et al. (2015) compared the quality of work between those working from the office to those working from home, the work was making sales calls via telephone. Similarly, Choudhury et al. (2021) compared remote patent examiners working from home to those working from anywhere, but in both cases the work entailed examining documents. We, however, compare work (auditing) conducted in-person onsite—with onsite employees sharing information and providing site tours—to work conducted remotely—with auditors working offsite and engaging with onsite employees and with their own team members only via audio/video technology. We therefore compare a particular form of remote work—in which workers and customers are not co-located—to the in-person scenario in which they are co-located. Such a setting provides greater variation in coordination and information flow, which we theorize will affect work quality.

2.2. Monitoring Business Partners' Management Practices

Our work also relates to the literature on the effectiveness of firms monitoring how well suppliers and other business partners comply with standards. Some buyers, for example, conduct monitoring to mitigate the risk of negative reputation spillovers, as when the media report that a fast food franchisee's customers were made sick by contaminated food preparation practices or that an apparel brand's supplier used child labor. Our work is most closely related to studies that identify factors that cause problems to go underreported (e.g., Gul et al. 2013, Short et al. 2016), which has largely been attributed to characteristics of individual auditors, auditing firms, audit arrangements, and institutional factors.

For example, audits are of lower quality when conducted (a) by auditors who are less-trained and less-experienced (Macher et al. 2011, Short et al. 2016) or have no experience auditing the site (Short et al. 2016, Ball et al. 2017), (b) by all-male audit teams (Short et al. 2016), or (c) later in an auditor's workday, when fatigue is more likely (Ibanez and Toffel 2020). Several audit firm characteristics have been found to erode monitoring quality; firms perform worse when they (a) have opportunities to cross-sell other services to those they are monitoring (Pierce and Toffel 2013), (b) face greater competition (Bennett et al. 2013), or (c) have performed the site's prior audit (Ibanez et al. 2023). A key arrangement associated with worse audit quality is when the audit is paid for directly by the monitored entity rather than by another interested party (Duflo et al. 2013, Short et al. 2016). These studies that uncovered various forms of bias associated with in-person auditing and inspecting concerned themselves, as do we, with auditors or inspectors underreporting results. We refer to such underreporting as poor audit quality, applying this concept from the financial auditing literature (e.g., DeAngelo 1981, Ronen 2010, Abbott et al. 2016).[4]

These studies have all focused on audits performed in-person onsite. We examine whether and how audit quality varies between such audits and those performed remotely and how such differences are moderated by key attributes of individual auditors and their firms.

3. Theory

Management systems standard certification audits assess the extent to which an organization is complying with the practices set forth in a management systems standard. Such standards focus on areas such as quality management (ISO 9001), environmental management (ISO 14001), and occupational health and safety (ISO 45001 and OHSAS 18001). Each standard has clauses and subclauses (henceforth "clauses") that address different areas of best practices.[5] Companies may voluntarily adopt these standards and hire an auditor to certify their adherence for several reasons, including a desire to adopt best policies and procedures and a desire or need to satisfy buyers or other stakeholders.

---

[4] For example, DeAngelo (1981) defines the quality of financial auditing as a function of the probability that the auditors discover and report problems in their client's accounting process; this concept has been widely used in the financial auditing literature (e.g., Ronen 2010) and also applied to internal auditing (e.g., Abbott et al. 2016).
[5] For example, the ISO 9001 standard has clauses for best practices for leadership, operations, and performance evaluation.

Auditors gather data to assess compliance with each clause by reviewing documents (such as policies, procedures, and training records), interviewing employees and managers, and directly observing processes and work activities (ISO 19011 Guidelines for Auditing Management Systems: Section 6.4.7). Instances of noncompliance are referred to as non-conformities or violations.

3.1. Quality Concerns of Remote Auditing

Service operations entail a coproduction process between service workers and customers, clients, patients, and so on (Vargo and Lusch 2004, Sampson and Froehle 2006). While remote work can reduce costs and travel time and enable more efficient planning, remote work that physically separates workers from those coproducers tends to reduce the availability of critical information needed for such coproduction, including access to diagnostic information ranging from patient symptoms to factory conditions. The result can be less-comprehensive—which we define as lower-quality—work. In our context, this separation can keep auditors from accessing information necessary to identify some violations at a site. Below, we argue that it is harder for remote auditors than for in-person auditors to (a) directly observe and detect some types of violation and (b) coordinate among audit team members, both of which lead to our first hypothesis:

H1. Remote audits will be of lower quality than in-person audits.

3.1.1. Information access challenges. The extent to which conducting diagnostic work remotely impedes quality will vary with the data-gathering techniques used. Assessing compliance with some management standards clauses requires auditors to review documents—a similar process whether conducted remotely or in-person. Other types of information are harder to access remotely. Technology-mediated interactions tend to reduce participant engagement and important nonverbal communication (Reid and Reid 2005, Bohannon et al. 2013).

Observing contextual attributes—which can include workers, equipment, procedures, and work settings—increases the information that can be incorporated into assessments and decisions (Tyre and von

Hippel 1997). In our context, auditors often rely on contextual knowledge gained at the audit site—by looking, listening, touching, and smelling—to better understand its operations and vulnerabilities and to assess compliance. Such on-the-spot observation can prompt ad hoc inquiries that can uncover violations. Site tours conducted remotely are limited by the audio-video technology—seldom as good as human senses—and to some extent by what the onsite employees choose to show. A recent survey found EHS professionals concerned that remote site tours risk resulting in assessments "limited to what those at the facility allow the auditor to see" and that some areas "may be overlooked…due to restriction of electronic equipment" in those areas (NAEM 2021: 19). These factors lead us to predict that a decline in audit quality associated with remote auditing will be greater for standards clauses for which compliance is assessed via direct observation than for clauses assessed via document review.

H2. The performance gap between remote and in-person audits will be wider (worse) for domains in which compliance is assessed via direct observation rather than document review.

3.1.2. Coordination challenges. Diagnostic service operations conducted remotely, such as remote auditing, entail physical separation not only between the remote workers and the subjects of their work, but also between team members conducting the work. Here, again, we theorize that separation reduces information flow between audit team members (virtual teams), reducing audit quality. Research has found that, compared to those working on physically separated teams, those working on in-person teams tend to (a) share a greater sense of commitment that leads them to spend more time on tasks and to work with more urgency and (b) engage in more information exchange that leads to them to adjust their tasks (Kiesler and Cummings 2002).

Virtual teams face coordination problems because physical separation makes it harder for members to gain a sense of shared knowledge (Cramton 2001). Uneven information distribution—or a lack of mutual knowledge—makes it harder for them to identify and resolve problems. One key mechanism identified in the literature is the lack of opportunities for spontaneous communication among physically distributed team members. Hinds and Mortensen (2005) find that such spontaneous communication enables team members to informally observe which tasks their colleagues are working on and/or struggling with. The authors theorize that virtual team performance suffers from a loss of spontaneous communication. Virtual teams are more likely to report less team trust (Peñarroja et al. 2013) and psychological safety (Gibson and Gibbs 2006), both of which can make it harder to openly communicate and resolve conflicts. Moreover, Yang et al. (2022) find that the shift to remote work has led workers to (a) engage in more asynchronous communication and less synchronous communication and (b) rely more heavily on tools—such as email and messaging—that exclude nonverbal communication. Lin et al. (2023) attribute their finding that remote teams of researchers produced fewer breakthrough ideas than in-person teams to the greater difficulties such teams face in sharing and assembling complex ideas.

In our setting, audits may be conducted by a single auditor or by a team. Auditor teams require frequent communication to exchange information and coordinate activities. The rules governing auditors of management systems standards require them to "confer periodically to exchange information, assess audit progress, and reassign work between the audit team members as needed" (ISO 19011: Section 6.4.4). These dependencies make auditing in a team a form of tightly coupled work, which is particularly challenging for remote teams (Olson and Olson 2000).

According to our interviews with the auditing company that provided our data, in-person audit teams tend to meet with one another on both a scheduled and unscheduled basis during an audit to discuss their observations and further investigative steps to better identify potential violations. This allows for a dynamic auditing process. Our interviews also revealed that, during remote team audits, auditors instead rely mostly on pre-scheduled meetings because their lack of co-location makes unplanned meetings harder—with fewer opportunities to coordinate and share best practices. We argue that these constraints impede information flow and coordination among remote audit team members, which leads to especially low-quality audits.

H3. The performance gap between remote and in-person audits will be wider (worse) for larger audit teams.

3.2. Mitigating Remote Auditing Quality Concerns via Prior Site Exposure

For financial auditors, familiarity with the site has long been thought to reduce in-person scrutiny; some regulators even require or recommend rotation of lead auditors or audit firms after a maximum number of years.[6] Research on in-person supplier auditing and regulatory inspections has found that auditors assigned to sites they have previously audited tend to conduct less-comprehensive audits (e.g., Short et al. 2016, Ball et al. 2017). One mechanism potentially driving this effect is "bounded awareness" (Chugh and Bazerman 2007), a cognitive bias that, in our context, suggests that auditors returning to the same site would overly focus on the problem areas they had identified in prior audits. Kumar and Chakrabarti (2012) argue that decision makers over-rely on information that supports their tacit knowledge gained from prior experiences, limiting their ability to recognize new problems.

We apply these arguments and findings, which focus on how in-person audit quality varies based on auditors' prior exposure to the site, to how remote audit quality might vary based on those auditors' prior in-person exposure to the site. Here, an additional factor comes into play: memory of the site layout and work practices observed in-person along with tacit knowledge of the site's vulnerabilities that can lead returning auditors to ask more questions that turn up more violations. Thus, having more in-person auditing experience at the site can help overcome the information access challenges of remote audits. We therefore expect that auditors with more in-person experience at the audited site will conduct higher-quality remote audits of that site than remote auditors without that experience.

> H4. The performance gap between remote- and in-person audits will be narrower for auditors with more in-person auditing experience at the site.

---

[6] For example, the Sarbanes-Oxley Act of 2002 requires US public companies to rotate lead audit partners every five years and the 2014 European Union (EU) Audit Regulation requires EU member states to create regulations mandating audit firm rotation every 10 years or sooner.

4. Data and Measures

4.1. Empirical Context

We obtained data from a company that provides auditing, certification, training, and consulting to organizations around the world and that required anonymity as a condition of sharing data with us. The company's auditing and certification practices are accredited by many leading accreditation bodies, which provides third-party assurance that its auditing meets international standards.

This company shared all certification audits it conducted from January 1, 2019 to October 27, 2021 pertaining to six of the most widely used management systems standards, covering management of quality (ISO 9001), environment (ISO 14001), information security (ISO 27001), occupational health and safety (OHSAS 18001 and ISO 45001), and medical device quality (ISO 13485). We begin the sample period in 2019 because our data provider began conducting a large portion of their audits remotely in 2020. But we also obtained data on the experience, training, and background of auditors starting in 2016 to avoid left- censoring when creating variables to measure these constructs.

In the course of obtaining and retaining certification, a site receives an *initial audit* to assess whether its management system meets the requirements of the given standard; if so, the audit firm issues a three-year certification.[7] Certified sites are subsequently subjected to *surveillance audits*, typically one a year. Auditors follow the same audit process in surveillance audits, but check only a subset of areas within the standard. In the third year, a *recertification audit* is conducted to provide another comprehensive assessment of the site's compliance; a successful audit results in a three-year renewal of certification. In our analysis, we focus only on surveillance audits due to their prevalence in the data and resulting statistical power; our dataset includes multiple surveillance audits for a given site being audited

[7] The initial audit is performed in two stages. In stage 1, auditors check whether the necessary procedures have been developed to meet the standard's quality benchmark. Stage 2 assesses whether those procedures were implemented and assesses the efficacy of their implementation. Failure to implement those procedures would result in no certification until the site provides evidence that the failures have been rectified.

to a particular standard—in many cases, some onsite and others remote.[8] While the COVID-19 pandemic increased the popularity of remote auditing, our dataset includes the hundreds of in-person surveillance audits that were conducted during every month of our sample period.

The raw data we received that addressed the six standards mentioned above included 125,727 surveillance audits of 48,956 sites during the nearly three-year period of our sample. Because some sites adopted more than one standard, this initial dataset corresponded to 63,569 site–standard dyads, which refers to all unique pairings of a given site with audits of a particular management systems standard (e.g., all of a particular site's ISO 9001 audits). We summarize our sample construction steps in Table A-1 in the Online Appendix. We omitted the 6,657 audits that had data anomalies that either resulted from data entry errors or did not reflect the typical course of audits according to our data provider.[9] We then omitted the 28,420 audits of sites that diverged from the standard practice of no more than one surveillance audit per year and instead had two or more; these more frequent audits tend to have a reduced scope compared to annual surveillance audits, according to the auditing company.[10] This left 90,650 surveillance audits in our sample. Using Poisson regression with site-standard fixed effects to estimate our model (described below) leads to the regression model mechanically dropping 55,403 audits that were either (a) a singleton site–standard panel (i.e., a given standard at a particular site had only one audit during our sample period) or (b) members of site–standard panels of which every surveillance audit reported zero violations, leaving

---

[8] Because initial audits are typically conducted once per site–standard, we cannot compare in-person to remote initial audits of a given site–standard. Because recertification audits are conducted roughly every three years and our sample only extends three years, our dataset has very few site–standard dyads with more than one recertification audit (and only a handful with at least one conducted in-person and at least one conducted remotely), which prevents us from analyzing recertification audits. We therefore exclude initial audits and recertification audits from our analysis, leaving their investigation to future research.

[9] Specifically, we omitted surveillance audits that (a) were missing a certificate number or auditor identifier, (b) prompted a change in certificate number, (c) were duplicates (same audit-format–site–standard–date), or (d) were associated with a site whose audits were conducted in an atypical sequence of initial, surveillance, and recertification audits or were missing audits from the typical sequence.

[10] We considered combining these multiple surveillance audits conducted in a calendar year for a single site–standard, thinking they might simply represent a surveillance audit split into parts. But our preliminary analysis led us to conclude that aggregated surveillance audits are not comparable to typical annual surveillance audits: the former required substantially more auditor-days than the latter, even after accounting for sites' employment level and industry, the factors that audit standards specify was determining the number of auditor-days to be allocated to each audit. We therefore dropped all site–standards that included multiple surveillance audits in a calendar year.

an estimation sample of 35,247 surveillance audits of 16,986 site–standards at 14,615 sites. (As described below, we conducted robustness tests that re-estimated alternative versions of our models on the 90,650 audits that include the 55,403 audits that were mechanically dropped by the fixed-effects Poisson model.)

Of the 35,247 surveillance audits in our estimation sample, 54% assessed compliance to the ISO 9001 quality management standard, 18% to the ISO 14001 environmental management standard, 12% to the ISO 27001 information security management standard, 11% to the OHSAS 18001 or ISO 45001 occupational health and safety management standards, and 4% to the ISO 13485 medical device quality management standard. While these audits were conducted in 109 countries, most took place in the United Kingdom (38% of sites), China (9%), India (7%), and the United States (6%).

Auditors assess a site's adherence to standards by reviewing documents, interviewing employees, and touring facilities. Audits have traditionally been performed onsite, but a growing number have been remote, with the use of remote audits accelerating during the COVID-19 pandemic to reduce exposure and travel.[11] According to our data provider, remote audits were performed from auditors' homes or offices, using smartphone cameras as the audio/video technology.[12] When remote audits were conducted by teams, there was no expectation for them to be collocated, and our data provider indicated that team members tended to work from their own homes or were split between home and the office. Remote audits and in-person audits are intended to entail the same level of scrutiny and thus to result in equal audit quality. According to our data provider, when auditors are assigned to an audit (whether in-person or remote), they are expected to start and finish that audit before beginning an audit for another client.

We follow other empirical studies of auditor bias by controlling for factors that might be correlated with violations and, assuming all else equal, considering an establishment's audits that report fewer violations to be of lower quality than those that report more violations. This approach has been used

[11] Starting in March 2020, audit format (remote or in-person) was determined by a risk assessment to assess whether an auditor could safely access the site in-person or whether it could be successfully audited remotely.
[12] Other audio/video technologies, such as virtual reality and smart glasses, were neither used by the audit company during our sample period nor widely used by other companies conducting internal environments, health, and safety audits in 2021, according to NAEM (2021).

in contexts including supplier codes of conduct (Short et al. 2016), vehicle inspections (Bennett et al. 2013, Pierce and Toffel 2013), and government inspections (Braithwaite and Makkai 1991, Gray and Shadbegian 2005, Duflo et al. 2013, Ibanez and Toffel 2020).

Auditors collect evidence to prove both compliance and noncompliance. When they do not find sufficient evidence of compliance with a clause, the site has a second chance to furnish it, mitigating the risk of reporting false violations. Specifically, management system auditing rules require auditors to review preliminary findings of noncompliance "with the auditee in order to obtain acknowledgement that the audit evidence is accurate and that the nonconformities are understood. Every attempt should be made to resolve any diverging opinions concerning the audit evidence or findings" (ISO 19011: Section 6.4.8). As a result, lower audit quality refers to auditors underreporting—not overreporting—violations.

4.2. Dependent Variables

We measured the degree to which a site does not comply with all clauses of a management systems standard as *violations*—the number of violations discovered and recorded in a surveillance audit. For example, violations identified during an audit of the ISO 9001 Quality Management Systems standard could correspond to the standard's clauses addressing site operations, leadership, or personnel; the site might not have established clear communication channels with clients or customers or devoted resources to establish proper quality assessment processes. This approach has been used in contexts including supplier codes of conduct (Short et al. 2016), restaurant inspections (Ibanez and Toffel 2020), and government health and safety inspections (Braithwaite and Makkai 1991, Gray and Shadbegian 2005).[13]

Discussions with the company that provided our data indicated that auditors tend to rely on one of several methods—document review, direct observation, and interviews—to identify noncompliance. For example, the ISO 9001 Quality Management Systems standard's clause 6.2 requires organizations to maintain documented information about their management system quality objectives and how to achieve

[13] Each violation is categorized as major if it impedes a management system's capability to achieve a site's objectives (e.g., financial success or product quality) or else as minor. In our sample, 97.5% of violations are minor. Below, we describe robustness tests finding that our results are nearly identical if we exclude major violations.

them, which auditors verify by reviewing those documents. In contrast, ISO 9001 clause 8.5.1 requires that suitable infrastructure and environment be used for site operations, which auditors tend to verify by directly observing the facility. Based on information obtained from our data provider for five of the six standards in our study (all except ISO 13485), we categorized each audit's *violations* into those associated with standard clauses for which violations were primarily detected by document review, by direct observation, or by employee interviews. Of the clauses in these five standards, 301 are primarily detected by document review (*document-review–based clauses*), 45 by direct observation (*direct-observation–based clauses*), and 61 by interviews (*interview-based clauses*), while 47 lacked a single primary detection mode (e.g., two modes were equally relied on).[14]

In our analysis of how quality concerns about remote auditing vary by violation-detection mode, we compare violations of direct-observation–based clauses to violations of document-review–based clauses, leaving aside interview-based violations as an intermediate category less affected by remote auditing. (We include interview-based clauses in a robustness test described below.)

4.3. Independent and Moderator Variables

Remote audit is a binary variable coded 1 when an audit was conducted remotely and 0 when conducted in-person. To develop a measure of audit team size, we observed that 80% of audits were conducted by a single auditor, 15% by a two-auditor teams, 4% by three auditors, and 1% by teams of four to nine. To capture the primary variation in our data and avoid results driven by outliers, we measured audit team size

---

[14] Examples of standards clauses primarily assessed via document review include requirements to establish a quality policy that is appropriate to the organization's purpose, context, and strategic direction and that includes a commitment to continual improvement of the quality management system (ISO 9001: 2015 clause 5.2.1 "Establishing the quality policy") and requirements to "keep documented information" showing that the organization has met its information security processes (ISO 27001: 2013 clause 8.1 "Inventory of assets"). Standards clauses primarily assessed via direct observation include requirements for the site to have implemented production and service provision under conditions consistent with its documented control procedures (ISO 9001: 2015 clause 8.5.1 "Control of production and service provision"). Standards clauses primarily assessed via interviews include requirements for the organization to monitor customers' perceptions of the extent to which their needs and expectations are being met (ISO 9001: 2015 clause 9.1.2 "Customer satisfaction") and to monitor and evaluate the performance and effectiveness of the management system in a manner that ensures comparable and reproducible results (ISO 27001: 2013 clause 9.1 "Monitoring, measurement, analysis and evaluation").

using a binary variable, *audit team,* coded 1 for audits conducted by more than one auditor and 0 for audits conducted by one auditor. (In unreported results, our results are robust to measuring team size as the actual number of auditors who conducted each audit.)

To measure an audit team's familiarity with the site from previous visits, we coded *average prior in-person site exposure (log)* by taking the sum of prior in-person audits any audit team member had conducted there from 2016 until the focal audit, dividing by the audit team size, and then taking the log of the result to reduce skew.

4.4. Control Variables

We control for several audit-, auditor-, and site-level variables that may be correlated with our outcome measure. We measured a site's complexity as the number of *staff-days (log)* scheduled to perform an audit. Staff-days is determined using International Accreditation Forum (IAF) documentation and is based on site's number of personnel and its complexity, including the number of buildings, applicable regulations, and process complexity.[15] We use the natural log to reduce skew.

Prior remote site exposure is a binary variable coded 1 when at least one audit team member had previously conducted a remote audit of the site since 2016, and 0 otherwise.[16] *Focal standard advanced training* is a binary variable coded 1 if any audit team member other than the audit team leader held a "lead assessor" qualification for the focal audit's standard anytime since 2016, and 0 otherwise.[17] *Maximum auditing experience (log)* is the maximum number of audits any auditor on the team had conducted since 2016 until the start of the focal audit, logged (after adding 1 to accommodate zeros) to

[15] Staff days for surveillance audits is based on the International Accreditation Forum document "Determination of audit time of quality, environmental, and occupational health & safety management systems" (IAF MD 5:2019), which states: "The starting point for determining audit time of management systems shall be identified based on the effective number of personnel, then adjusted for…significant factors" (Section 3.7), including the site's logistical complexity, multilingual staff, degree of industry regulation, and process complexity (Section 8). The IAF document's Annex A specifies how audit staff time is to be adjusted based on the site's number of personnel (Table QMS 1) and on various dimensions of the site's complexity (Figure QMS 1).

[16] We chose this measure instead of measuring *average prior remote site visits (log)* because the binary variable captures 87% of variation in *average prior remote site visits (log)*.

[17] We recoded this variable to 0 for the 3,912 audits for which this information was unavailable and created a dummy variable coded 1 to flag those observations and 0 otherwise.

reduce skew.[18] *Percent outsourced* is the percentage of audit team members who were contract employees rather than employees of the auditing company.

Multi-standard audit is a binary variable coded 1 if an audit for a different standard was conducted at the same time as the focal audit, and 0 otherwise. We created audit-year fixed effects to control for secular factors and audit-sequence fixed effects to reflect how many audits a site received from the audit company before the focal audit (1st audit, 2nd audit, etc.).[19]

Summary statistics and correlations are reported in Tables 1 and 2.

5. Empirical Models and Results

5.1. Empirical Specification

We estimate four models to test our theory. For Models 1, 3, and 4—testing H1, H3, and H4, respectively—the unit of observation is an audit conducted to assess the implementation of a given standard at a given site on a given date. For Model 2, which tests H2, we divide each audit into two observation that are identical except that in one, *direct-observation–based clause* is coded 1 and the dependent variable *violations* includes the sum of violations of standards clauses whose primary detection mode is direct observation, while in the other, *direct-observation–based clause* is coded 0 and the dependent variable *violations* includes the sum of violations of standards clauses whose primary detection mode is document review. In sum, the unit of observation for Model 2 is an audit's violation-detection mode.

Model 1 predicts *violations* as a function of *remote audit* and other controls (described below) to test the main effect of remote auditing on audit quality. Model 2 also includes an interaction term between *remote audit* and *direct-observation–based clauses* to assess whether quality concerns with regard to remote audits are more severe for management standard clauses whose assessment relies on auditors'

---

[18] As described below, a robustness test that replaced *maximum auditing experience (log)* with the audit team's *minimum auditing experience (log)* yielded nearly identical results.
[19] Before creating audit-sequence fixed effects, we winsorized the sequence count at the 95th percentile (16th audit) to minimize the influence of outliers.

direct observation as opposed to document review. Model 3 instead includes an interaction term between *remote audit* and *audit team* to assess whether quality concerns with regard to remote audits are more severe for audits conducted by teams than for those conducted by a single auditor. In Model 4, we include an interaction term between *remote audit* and *average prior in-person site exposure (log)* to assess whether quality concerns with regard to remote audits are attenuated when the remote audit team includes anyone who has conducted in-person audits of the site before.

Our models include the following control variables. Site-standard–dyad fixed effects (e.g., Site A–ISO 9001) control for time-invariant factors (observed and unobserved) associated with each combination of site and standard that could affect compliance, such as industry and location.[20] We include *staff-days (log)* because more time at the site provides more opportunity to uncover violations. We control for *audit team* because more auditors assigned to an audit might result in more violations being cited. To account for the possibility that an auditor's advanced qualifications for the audit's standard could affect audit quality, we include *focal standard advanced training*.

To control for the possible effect of audit team members' familiarity with the site on their findings, we include *average prior in-person site exposure (log)* and *prior remote site exposure*. Since cumulative experience influences audit outcomes (Macher et al. 2011, Short et al. 2016, Ball et al. 2017), we include *maximum auditing experience (log)*. We include *percent outsourced* because different audit firms provide different training and have been shown to differ in the number of violations they cite (Ibanez et al. 2024). *Multi-standard audit* accounts for the possibility that sites' violation counts could be affected by their being audited against several standards simultaneously.

We include audit-sequence fixed effects because sites might become more compliant the more they are audited, having had more opportunities to learn about and resolve violations. All models include audit month-year fixed effects (January 2019, February 2019, and so on through October 2021) to account

[20] As described below, Model 2's results are robust to an alternative specification that includes site–standard–detection-mode fixed effects.

for secular factors that might affect auditor stringency and/or violation counts, including the COVID-19 pandemic and related policies that led some sites to operate at reduced capacity with fewer workers.[21]

5.2. Identification Strategy

5.2.1. Remote audit assignment. If our results are to provide an unbiased estimate of the effect on violation counts of conducting audits remotely as opposed to in-person, unobserved variables should not be correlated with both an audit's outcome and the decision to perform the audit remotely. Our data provider indicated that it first adopted remote auditing to reduce costs, and that it gained much wider use with the COVID-19 pandemic. The decision to conduct an audit remotely versus in-person was a two-step process. First, when the audited entity's country regulations allowed in-person auditing, the auditing company used its standardized "Pre-visit COVID-19 checklist" to determine whether its auditors could do so safely.[22] Second, the company assessed whether the site had the IT infrastructure to support a remote audit. Any site that could not be visited in-person safely and did not have the infrastructure for a remote audit would not be audited until it was safe to do so in-person. Moreover, audited sites would not receive a remote audit if they did not agree to it.

We hypothesize that remote audits will be of lower quality than in-person audits, as evidenced by reporting fewer violations. To ensure that such a result is not driven by decisions to conduct audits remotely (versus in-person) when site–standards were expected to have fewer violations, we estimated a regression model that predicted whether the focal audit was conducted remotely (*remote audit*) based on the site–standard's prior audit's violation count and all other independent and control variables used in Model 1. A negative coefficient on the prior audit's violations would indicate that "cleaner" establishments were more likely to be assigned remote audits, which risks confounding whether a finding

---

[21] As described below (and reported in Table A-4 of the Online Appendix), results are robust to controlling for geographically varying secular effects that include interactions between month-year fixed effects and country fixed effects.

[22] Items in this checklist include the number of positive COVID-19 cases among site employees in the past weeks, whether the site has implemented social distancing guidelines in all areas, and whether there is a dedicated area where auditors can maintain physical distance from others.

that remote audits indeed yielded fewer violations really did result from lower-quality auditing (as we hypothesize in H1) or from sites prone to fewer violations being more likely to be assigned to remote audits. In fact, estimating this model that predicted *remote audit* using logistic regression or OLS regression yielded a coefficient on the prior audit's violations that was *positive* and marginally statistically significant, (Model 1 using logistic regression: $\beta = 0.058$; $p < 0.10$) or not statistically significant (Model 2 using OLS regression: $\beta = 0.005$; $p = 0.31$), yielding no evidence that "cleaner" sites were assigned to remote audits. We report these results in Table A-2 in the Online Appendix.

5.2.2. Auditor assignment. There are three reasons we believe the assignment of auditors is plausibly exogenous to our focal variables and violation counts, conditional on the variables included in our models. First, we learned through conversations with the audit company that auditors qualified to audit a given site's industry were assigned to audits (of any format) based on their availability. (Auditors also needed industry-specific qualifications for sites in some industries, such as aerospace and medical devices). This provides no evidence that the auditor assignment process might bias our results; for example, there is no evidence that "more stringent" auditors were disproportionately assigned to in-person audits. Moreover, including site-standard fixed effects in our models means that we compare in-person to remote audits of the same site–standard over time (e.g., Factory ABC's adherence to the ISO 9001 standard) and the site's industry is exceedingly unlikely to change during our limited sample period.

Second, including site-standard fixed effects holds constant the minimum auditing qualifications needed to audit a particular site for a particular standard, which mitigates the possible concern that arises in a cross-site analysis that sites in more complex industries require both in-person audits (due to their complexity) and audit teams with more qualifications that might enable them to find more violations.

Third, we learned that while the audit company first decided whom to assign to a particular audit and then whether the audit would be in-person or remote, the second decision process (as described above) took no account of which auditors had been assigned. This further mitigates concerns that individual auditor attributes might influence their assignment to an audit or how it was conducted.

We explored whether our primary results might be driven the assignment of some auditors to conduct only remote audits and others to conduct only in-person audits. First, we note that 98% (34,660 of 35,247) of the audits in our primary sample were conducted by at least one auditor who conducted *both* in-person and remote surveillance audits during our sample period. Thus, almost no audit teams were comprised of only those who exclusively conducted remote or in-person audits. Second, we estimated our models on the subsample that excluded these 2% of audits, with results (unreported) nearly identical to those estimated from our full sample.

5.3. Results

Given that our dependent variable (*violations*) is a count, we estimate our models using a Poisson pseudo-maximum likelihood fixed-effects regression.[23] We report standard errors clustered at the audited site. We interpret effect sizes in terms of incidence rate ratios (IRR) and average marginal effects (AME).

Table 3 reports our regression results. Before addressing the results concerning our theorized relationships, the statistically significant effects of several control variables merit discussion. In Model 1, the positive coefficient on *staff-days (log)* indicates that audits with more staff-days allotted tend to yield more violations, possibly because the audited sites are larger (in terms of employment) and/or their management systems are more complex, offering more opportunity to find violations. Similarly, the negative coefficient on *percent outsourced* means that audit teams with a larger proportion of auditors who are not direct employees of the audit provider report fewer violations, consistent with Ibanez et al. (2024).

Model 1 tests the overall effect of remote work on audit quality. The negative and statistically significant coefficient on *remote audit* ($\beta = -0.294$, $p < 0.01$, IRR = 0.75, AME = -0.40) indicates that remote audits yield an average of 25% fewer violations than in-person audits—1.17 versus 1.58—which supports H1.

[23] Poisson panel estimators are consistent even when data are not distributed, provided the conditional mean is correctly specified, and impose weaker distributional assumptions than negative binomial regression (Cameron and Trivedi 2010).

The results of Model 2 indicate that audit quality is worse for violations of clauses for which auditors primarily rely on direct observation than for the omitted category, for which they primarily rely on document review, based on the statistically significant negative coefficient on *direct-observation–based clauses* ($\beta = -1.722$, $p < 0.01$, IRR = 0.18).[24] For clauses assessed primarily via document review (the omitted detection-mode category), remote audits yield significantly fewer violations than in-person audits, as indicated by the statistically significant negative coefficient on the main effect of *remote audit* ($\beta = -0.210$, $p < 0.01$, IRR = 0.81). One possible explanation is that compliance areas primarily audited through document review also use interviews or direct observation to confirm the violations and, in remote audits, the increased difficulty in obtaining information through these secondary methods makes it harder to gather enough such evidence. The statistically significant negative coefficient on the interaction term *remote audit × direct-observation–based clauses* ($\beta = -0.617$, $p < 0.01$, IRR = 0.54) indicates that the quality of remote audits is further eroded for standards clauses for which violation detection relies on direct observation, consistent with our theory and supporting H2.[25] In Figure 1, average marginal effects indicate that for clauses primarily detected via document review, remote audits report 21% fewer violations (0.96) on average than in-person audits (1.21). Assessment of clauses primarily detected via direct observation suffer an even greater "remote penalty," with remote audits reporting an average of 0.09 such violations compared to 0.22 for in-person audits—a 59% decline.

Model 3 tests whether the decrement in remote audits' violations (compared to in-person audits) is larger for auditor teams than for solo auditors. The statistically significant negative coefficient on *remote audit × audit team* ($\beta = -0.299$, $p < 0.01$, IRR = 0.74) indicates that it is, which supports H3; the quality decrement of remote audits conducted by audit teams is greater than the decrement for audits

---

[24] Recall that for Model 2—unlike the other models reported in Table 3—the unit of analysis is an audit at the site–standard–detection-mode level, in which each audit is represented by two observations, one in which *violations* is the sum of non-compliance instances of clauses whose primary detection mode is direct observation and one in which *violations* is the sum of non-compliance instances of clauses whose primary detection mode is document review.

[25] As described below, this result is robust to an alternative model that includes site–standard–detection-mode fixed effects (e.g., Site A–ISO 9001–direct-observation–based) instead of site-standard fixed effects.

conducted by a single auditor (*remote audit* β = -0.244, p < 0.01, IRR = 0.78). Figure 2 reports average marginal effects: for audits conducted by a single auditor, remote audits averaged 1.20 violations, 22% lower than the 1.53 average for in-person audits. For team audits, remote audits averaged 1.02 violations, 42% lower than the 1.75 average for in-person audits. Put another way: for in-person auditing, audit teams report more violations (1.75) than solo auditors, but for remote auditing, audit teams report fewer violations (1.01) than solo auditors (1.20). These findings are consistent with our second proposed mechanism: remote audit teams' coordination problems. In in-person settings, meetings amongst team members tend to occur more frequently and at more natural points throughout the audit due to the auditors' co-location. In remote audits, lack of co-location may reduce the frequency and natural timing of these meetings, which undermines work quality.

Model 4 reports results of testing whether the quality decrement of remote auditing is attenuated for auditors who had conducted more in-person audits at the site. We begin by interpreting the negative coefficient on the main effect of *average prior in-person site exposure (log)* (β= -0.201, p < 0.01, IRR = 0.82) as indicating lower-quality in-person audits by teams with more in-person experience at the site, a result consistent with findings that auditors returning to a site report fewer violations than auditors new to the site (Short et al. 2016, Ball et al. 2017). This audit-quality decrement of site exposure is attenuated for remote auditing, as evidenced by the statistically significant positive coefficient on the interaction term *remote audit × average prior in-person site exposure (*β = 0.087, p < 0.01, IRR = 1.09).[26] The IRRs for these effects indicate that each one-unit increase in *average prior in-person site exposure (log)* is associated with a nearly 18% reduction in in-person audit violations (an "auditor familiarity penalty"), but this effect is attenuated by 9% for remote audits. The average marginal effects, displayed in Figure 3,

[26] Because the maximum value of unlogged *average prior in-person site exposure* is 48, one concern is that outlier values drive our results for this construct, even though we use its natural log to compress outliers. We therefore estimated a model that top-codes this variable at the 95th percentile (an average of 9 prior in-person site visits) before taking its natural log. Using this top-coded variable continues to yield a positive and statistically significant coefficient on the interaction term, consistent with our primary results and indicating that they are not driven by outlier values of *average prior in-person site exposure*.

indicate that, for example, tripling the (unlogged) average number of prior in-person audits from approximately two to six (that is, from 0.7 to 1.8 log points along the x-axis) reduces the average violation gap between remote and in-person audits by 0.2 violations, from 0.35 to 0.15.[27] Figure 3 indicates that the moderating effect of more prior in-person exposure on audit quality is driven by a steeper decline in in-person audit quality compared to the more modest decline in remote audit quality. These results indicate that prior in-person site exposure only partially attenuates quality concerns for remote auditing and—supporting H4—that greater in-person exposure does not increase remote audit quality but instead attenuates the "familiarity penalty" that motivates calls for auditor rotation.

5.4. Extension

Our main analysis revealed that auditors' prior in-person exposure to a site attenuates quality concerns for remote auditing. We now explore a second element that might also affect remote auditing quality: the extent to which the auditors' experience was concentrated on the management standard of the focal audit. A long literature has documented performance benefits from focusing on specific tasks (e.g., Taylor 1911, Fayol 1967); we explore whether this applies to remote auditing. To measure the extent to which the audit team included an auditor highly focused on the management system being audited, we calculated the proportion of each member's prior auditing experience that focused on that standard and assigned the maximum value to the team's *maximum management standard focus* for the audit. To estimate the influence of auditor focus on the quality of remote audits, we modify Model 1's specification by adding *maximum management standard focus* and its interaction with *remote audit.*[28] The results, reported as Model 5 in Table 3, yield a positive and statistically significant coefficient on the interaction term (β = 0.148, $p < 0.01$), indicating that *maximum management standard focus* attenuates the quality decrement

---

[27] We arrive at the estimation of 0.2 violations as follows. An increase from 2 to 6 prior onsite audit visits—or, taking the logs, from 0.7 to 1.8 log points of *average prior in-person site exposure (log)*—corresponds to a decline in *predicted average violations* of 0.34 for in-person audits (from roughly 1.73 to 1.38) and 0.15 for remote audits (from roughly 1.24 to 1.09). The difference-in-differences between 0.35 and 0.15 is 0.20.

[28] We recoded missing values of *maximum management standard focus* to 0 and added to this model a dummy variable coded 1 when we conducted this recoding and 0 otherwise.

of remote auditing. Accumulating auditing experience with a given standard appears to partially mitigate remote auditors' information access challenges. The small, nonsignificant coefficient on the main effect of *maximum management standard focus* yields no evidence that such focus improves in-person auditing. Figure 4 graphs the average marginal effects of these results; increases in *maximum management standard focus* are, on average, associated with an attenuated gap in violations reported by remote versus in-person audits, given a sharper increase in violations reported by remote audits than by in-person audits as *maximum management standard focus* increases.

5.5. Robustness Tests

Our results are robust to several alternative measures and samples.

Controls. First, to assess whether our results were driven by how we control for audit team experience, we re-estimated our models replacing *maximum auditing experience (log)* with *minimum auditing experience (log)*. These results, reported in Table A-3 in the Online Appendix, are nearly identical to our primary findings.

Second, we explored a more flexible approach to control for secular effects (including pandemic severity) that might have changed working conditions (and thus violations) and the tendency to conduct remote (versus in-person) audits. Specifically, we included a full set of interactions between *month-year fixed effects* and the sites' *country fixed effects.* This lets temporal factors vary by country, accounting for the possibility that the influence of the pandemic and other secular shocks on working conditions might have varied by country. The results, reported in Table A-4 in the Online Appendix, yield coefficients on our hypothesized variables very similar to our primary results.

Third, as a robustness test for Model 2, we replaced the site-standard fixed effects with site–standard–detection-mode fixed effects, which provides estimated effects on violations of remote auditing within a detection-mode (direct observation or document review) across audits of a given standard at a site. With this more conservative approach, the main effect of *direct-observation–based clauses* is absorbed by the fixed effects and observations are mechanically dropped when (a) they are the only one associated with a given site–standard–detection-mode or (b) there are multiple observations associated with a given site–standard–detection-mode but all report zero violations. This alternative specification continues to support H2 by yielding a statistically significant negative coefficient on the interaction term *remote audit* × *direct-observation–based clauses* (β = -0.408, $p < 0.01$, IRR = 0.66); results are reported in Table A-5 in the Online Appendix.

Fourth, we examined the implication of our data lacking auditor gender for nearly half of our sample, which led us to omit audit team gender mix from our main specification despite prior research finding that it can affect audit results (Short et al. 2016). To assess the extent to which our inability to control for this in our primary models might have biased the estimates of our hypothesized effects, we twice re-estimated each model from Table 3 on the subsample of audits for which we had data on the audit team's gender composition: once using our main specification and once adding the dummy variable *female on audit team*. As shown in Table A-6 in the Online Appendix, the positive coefficient on *female on audit team* indicates that having at least one woman on an audit team was associated with higher audit quality, consistent with prior research (Short et al. 2016).[29] More importantly for our purposes, the coefficients on the hypothesized variable in these models are nearly identical and Wald tests find none of the very minor differences statistically significant. This yields no evidence that our primary estimates in Table 3 are likely to be biased by not controlling for audit team gender composition.

Fifth, one might question whether our results are influenced by our Poisson regression models with site-standard fixed effects mechanically dropping the 32,568 audits that pertain to (a) singleton site–standard dyads or (b) site–standard panels whose every audit reports zero violations. To include these additional observations in our estimation sample, we re-estimated our models after (a) replacing site-standard fixed effects with random effects or (b) simply omitting these fixed effects. These models,

---

[29] The effect size is roughly comparable as well. For example, in Model 1b of Appendix Table A-6, the *female on audit team* coefficient (β = 0.094; SE = 0.043) indicates that teams with at least one woman reported 9% more violations, which is similar to the 5%–6% effect size that Short et al. (2016) found for global supply chain code-of-conduct auditing.

reported in Table A-7 in the Online Appendix, yield results estimated on this larger sample that are very similar to those of our primary models, confirming that our results are not driven by our fixed-effects models dropping these particular observations.

Dependent variable. We conducted three sets of robustness tests of whether our results were robust to alternative dependent variables. First, given that our analysis examines six standards that differ in the maximum number of violations that auditors report (and other time-invariant factors), our primary models control for these differences by including site-standard fixed effects.[30] To more explicitly account for the maximum number of violations cited for each standard, we estimated alternative versions of our models that predict ratios of these maximum values. Specifically, we re-estimated Models 1, 3, 4, and 5 using an alternative dependent variable that normalizes the audit's results by the standard's maximum number of violations observed in our sample—*violations as proportion of standard's maximum*—calculated by dividing *violations* by the sample's maximum number of violations for this standard. Similarly, for Model 2, the dependent variable is *violations as proportion of standard–detection-mode's maximum,* calculated by dividing *violations* by the sample maximum number of violations for this standard–detection-mode. Estimating these model specifications—otherwise akin to our primary models—yielded, as reported in Table A-8 in the Online Appendix, statistically significant hypothesized coefficients of the same sign as our primary models, indicating those models' robustness to this alternative specification.

We then re-estimated our five models in Table 5, this time predicting *minor violations*—97.5% of all violations in our sample—to see if our primary results might be driven by the rare major violations. The results, reported in Table A-9 in the Online Appendix, were nearly identical to our primary findings, assuaging this concern.

---

[30] Merely including standard fixed effects would account for these time-invariant differences between standards. Our approach is more flexible, allowing not only time-invariant differences between these standards, but also time-invariant factors associated with sites' implementation of each standard.

Finally, we assessed whether auditor discretion might be driving our primary findings. Auditors have discretion about whether to assess some standard clauses during a site's first or second surveillance audit in the three-year certification cycle. We refer to these as "discretionary clauses," in contrast to the "mandatory clauses" that must be audited in every surveillance audit. Our primary results could be biased if remote auditors chose to audit fewer or less–violation-prone discretionary clauses, which would mean that our finding that remote audits reported fewer violations might be attributable to that choice rather than to our posited mechanisms. We therefore constructed an alternative dependent variable—the number of violations associated with mandatory clauses—by omitting all audits of the ISO 13485 standard because we could not obtain a list of its mandatory clauses, and then considering only violations of the mandatory clauses of the remaining standards. These models, reported in Table A-10 in the Online Appendix, indicate that our primary results are robust to the omission of discretionary clauses.

Alternative samples. We conducted robustness tests by re-estimating our primary models on various subsamples. First, we re-estimated our model excluding establishments in the European Union because many of those countries adopted particularly aggressive COVID policies. These results, reported in Table A-11 in the Online Appendix, are very similar in magnitude and statistical significance to our primary results estimated on the entire sample.

Second, concerned that our results might be influenced by different types of site–standards having only in-person audits or only remote audits, we re-estimated our models on the subsample of site–standard dyads subjected to at least one in-person audit and at least one remote audit during our sample period. We also re-estimated our models on the subsample of audits conducted by teams that have both in-person and remote auditing experience during our sample, omitting the few audits by auditors who had conducted only remote or in-person audits. These results, reported in Tables A-12 and A-13, respectively, in the Online Appendix, continue to yield statistically significant support for all of our hypotheses, indicating that our results are robust to this concern.

Third, recalling that Model 2 tests the information access mechanism by comparing the effect of remote audits on finding violations primarily detected by direct observation to finding those primarily detected by document review, we also included in our sample a third record per audit where violations refers to those primarily detected by interviews and included in our specification a binary variable *interview-based clauses* coded 1 to denotes these observations) and its interaction with *remote audit*. Table A-14 in the Online Appendix shows a statistically significant negative coefficient on *remote audit × direct-observation–based clauses* ($\beta = -0.617$, $p < 0.01$) that reinforces our primary finding (Table 3, Model 2) that remote audit quality concerns are worse for clauses for which violation discovery relies on direct observation, even when the model includes interview-based audits. The nonsignificant (and slightly positive) coefficient on *remote audit × interview-based clauses* yields no evidence that remote auditing impedes the reporting of violations discovered primarily by interviews.

6. Discussion and Conclusion

We find that audit quality is lower when audits are conducted remotely and our analysis yields evidence consistent with our two theorized mechanisms: impediments to information access and to within-team coordination. Remote auditors appear to have problems gathering information in ways that proxy in-person auditors' direct observation during site tours, which we attribute to both their senses and their agency being hampered by technology mediation. We also find that remote audits conducted by multiple auditors suffer more quality degradation than those conducted by individual auditors, suggesting that the team collaboration necessary to process the gathered evidence is especially difficult with remote audits. Finally, we theorize and find evidence that an audit team with more experience conducting in-person audits of the site attenuates the quality decrement of a remote audit. Our empirical extension indicates that quality concerns for remote audits are somewhat mitigated when the audit team includes at least one auditor whose experience is concentrated on the standard being assessed.

Our findings contribute to the remote work literature by showing that work quality can suffer when working remotely changes not only where but how the work is conducted; in our case, by separating the workers both from the evidence they are assessing and from the coworkers with whom they need to collaborate. These contextual differences likely explain why we find that working remotely degrades quality while prior studies did not find such quality differences in remote work settings—answering customer calls and reviewing patent documents—that had neither of these attributes (Bloom et al. 2015, Choudhury et al. 2021). Given that many other work settings besides auditing—from telemedicine to drone-based crop monitoring to regulatory inspections—require direct observation, our results suggest that managers in those domains should take special care to control the quality of remote work.

The literature on the quality of business monitoring has focused on how the quality of in-person assessments such as quality auditing and regulatory inspections are affected by individual and organizational attributes (e.g., Macher et al. 2011, Bennett et al. 2013, Duflo et al. 2013, Pierce and Toffel 2013, Short et al. 2016) and by scheduling (e.g., Ibanez and Toffel 2020). We go beyond the in-person context to examine how audit quality is affected by remote work. Our findings are especially relevant to this literature stream given the increase in and expected persistence of remote audits not only within our context of management system certification (IAF, ILAC, and ISO 2021) but also in various forms of regulatory inspection (Mofid et al. 2021, US Department of Health and Human Services 2024). Our results also support concerns about the comprehensiveness of remote inspections (e.g., US Department of Labor 2021) and suggest caution to enthusiasts of remote auditing (e.g., IAF, ILAC, and ISO 2021). We lend nuance to the debate about remote work by identifying mechanisms that moderate quality concerns; these mechanisms suggest actions companies and regulators can take, such as avoiding larger remote auditing (or inspection) teams and developing better technologies and procedures for remote site tours. Moreover, our results can help companies and regulators design hybrid or blended auditing or inspection approaches to capture some of the benefits of remote auditing—such as increased flexibility, improved productivity, and reduced travel time and cost—while limiting its harm to audit quality. For example, hybrid schemes might pair remote document review with in-person direct observation or target fully in-person (that is, more comprehensive) audits to high-risk sites, fully remote audits to lower-risk sites, and hybrid audits to intermediate-risk sites.

While our results are based on the remote work context of auditing adherence to management systems standards, we suspect that they generalize to other contexts in which working remotely reduces or eliminates in-person information exchange. Examples include teaching, conducting medical exams, auditing working conditions at global supply chain factories, assessing restaurant franchisees' adherence to franchisors' procedural requirements, and government inspections that monitor compliance with occupational safety, environmental, and restaurant hygiene regulations.

Our study's limitations offer future research opportunities. The auditing company we studied conducted remote audits using smartphones as their sole audio/video technology for site tours, which was typical practice at the time (NAEM 2021). Because technologies such as virtual reality and smart glasses can provide higher-fidelity imagery, research should explore whether they improve remote auditing.

While our sample period includes the COVID-19 pandemic, the audit company continued to conduct hundreds of in-person and hundreds of remote audits of our focal six standards every month throughout that period. The pandemic might have affected site operations in ways that could alter compliance with management systems standards; for example, by reducing production volume in ways that reduced violations. But our use of month-year fixed effects and our within–site-standard empirical approach, along with our alternative approach using country x month-year fixed effects, accounts for such concerns. To bias our estimates of the effect of audits being conducted remotely versus in-person, the pandemic would need to differentially affect remote auditors' versus in-person auditors' propensity to detect violations.[31] While we have no evidence that this occurred, it calls for future research examining other time periods.

It is also possible that remote auditing yielded less-comprehensive audits due to auditors' increased mental exhaustion from doing a much larger portion of their work online. While our study cannot account for such potential mental costs of remote work, future research could. Finally, remote auditing might be a skill that takes time to acquire and the observed quality deficit might attenuate with

---

[31] For example, one might speculate that during the months when the pandemic was especially raging, auditors conducting remote audits might have been more distracted in ways that impaired their work quality than auditors conducting in-person audits (e.g., due to remote auditors' need to supervise their children who were being remotely schooled). However, it is also possible that auditors conducting in-person inspections might have been unusually distracted; for example, by worries about how the pandemic was affecting their friends and families.

experience. While our context of annual surveillance audits and our dataset's three-year sample period prevents us from assessing this, future research should.

References

Abbott LJ, Daugherty B, Parker S, Peters GF (2016) Internal audit quality and financial reporting quality: The joint importance of independence and competence. *Journal of Accounting Research* 54(1): 3–40.

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

Barrero JM, Bloom N, Davis SJ (2021) Why working from home will stick. NBER Working Paper No. 28731 (April).

Bavafa H, Hitt LM, Terwiesch C (2018) The impact of e-visits on visit frequencies and patient health: Evidence from primary care. *Management Science* 64(12): 5461–5480.

Bell DR, Gallino S, Moreno A (2018) Offline showrooms in omnichannel retail: Demand and operational benefits. *Management Science* 64(4): 1629–1651.

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

Bettinger EP, Fox, L, Loeb S, Taylor ES (2017) Virtual classrooms: How online college courses affect student success. *American Economic Review* 107(9): 2855–2875.

Bloom N, Liang J, Roberts J, Ying ZJ (2015) Does working from home work? Evidence from a Chinese experiment. *Quarterly Journal of Economics* 130(1): 165–218.

Bohannon LS, Herber AM, Pelz JB, Rantanen EM (2013) Eye contact and video-mediated communication: A review. *Displays* 34(2): 177–185.

Braithwaite J, Makkai T (1991) Testing an expected utility model of corporate deterrence. *Law & Society Review* 25(1): 7–40.

Buell RW, Campbell D, Frei FX (2010) Are self-service customers satisfied or stuck? *Production and Operations Management* 19(6): 679–697.

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

Campbell D, Frei F (2010) Cost structure, customer profitability, and retention implications of self-service distribution channels: Evidence from customer behavior in an online banking channel. *Management Science* 56(1): 4–24.

Choudhury PR, Foroughi C, Larson B (2021) Work-from-anywhere: The productivity effects of geographic flexibility. *Strategic Management Journal* 42(4): 655–683.

Chugh D, Bazerman MH (2007) Bounded awareness: What you fail to see can hurt you. *Mind & Society* 6(1): 1–18.

Cramton CD (2001) The mutual knowledge problem and its consequences for dispersed collaboration. *Organization Science* 12(3): 346–371.

DeAngelo LE (1981) Auditor size and audit quality. *Journal of Accounting and Economics* 3(3): 183–199.

Duflo E, Greenstone M, Pande R, Ryan N (2013) Truth-telling by third-party auditors and the response of polluting firms: Experimental evidence from India. *Quarterly Journal of Economics* 128(4): 1499–1545.

Fayol H (1967) *General and Industrial Management* (Pitman, London).

Gibson CB, Gibbs JL (2006) Unpacking the concept of virtuality: The effects of geographic dispersion, electronic dependence, dynamic structure, and national diversity on team innovation. *Administrative Science Quarterly* 51(3): 451–495.

Gray WB, Shadbegian R (2005) When and why do plants comply? Paper mills in the 1980s. *Law & Policy* 27(2): 238–261.
Gul FA, Wu D, Yang Z (2013). Do individual auditors affect audit quality? Evidence from archival data. *Accounting Review* 88(6): 1993–2023.
Hinds PJ, Mortensen M (2005) Understanding conflict in geographically distributed teams: The moderating effects of shared identity, shared context, and spontaneous communication. *Organization Science* 16(3): 290–307.
IAF, ILAC, and ISO (2021) Use of remote techniques supported by IAF/ILAC/ISO survey. Retrieved March 9, 2023, https://ilac.org/latest_ilac_news/use-of-remote-techniques-supported-by-iaf-ilac-iso-survey/.
Ibanez MR, Palmarozzo A, Short JL, Toffel MW (2024) Second- vs. third-party audit quality: Evidence from global supply chain monitoring. Working paper, Harvard Business School, Boston.
Ibanez MR, Toffel MW (2020) How scheduling can bias quality assessment: Evidence from food-safety inspections. *Management Science* 66(6): 2396–2416.
Internal Audit Foundation and AuditBoard (2021) The remote auditor: Challenges, opportunities, and new ways of working (Internal Audit Foundation and AuditBoard, Lake Mary, FL). https://www.theiia.org/en/content/research/foundation/2021/the-remote-auditor-challenges-opportunities-and-new-ways-of-working/.
Kiesler S, Cummings JN (2002) What do we know about proximity and distance in work groups? A legacy of research. Chapter 3 in P. J. Hinds, S. Kiesler, eds. *Distributed Work*, 57–80 (MIT Press, Cambridge, MA).
Kumar AJ, Chakrabarti A (2012). Bounded awareness and tacit knowledge: Revisiting Challenger disaster. *Journal of Knowledge Management* 16(6): 934–949.
Lin Y, Frey CB, Wu L (2023) Remote collaboration fuses fewer breakthrough ideas. *Nature* 623(7989): 987–991.
Macher JT, Mayo JW, Nickerson JA (2011) Regulator heterogeneity and endogenous efforts to close the information asymmetry gap. *Journal of Law & Economics* 54(1): 25–54.
Mofid S, Bolislis WR, Brading C, Hamilton N, Hardit C, Nagaoka M, Parain J, Zanta M, Kühler TC (2021) The utility of remote inspections during the COVID-19 health emergency and in the postpandemic setting. *Clinical Therapeutics* 43(12): 2046–2063.
NAEM (2021) The emergence of virtual EHS audits (National Association for Environmental, Health & Safety and Sustainability Management, Washington, DC).
Olson GM, Olson JS (2000). Distance matters. *Human-Computer Interaction* 15(2-3): 139–178
Peñarroja V, Orengo V, Zornoza A, Hernández A (2013) The effects of virtuality level on task-related collaborative behaviors: The mediating role of team trust. *Computers in Human Behavior* 29(3): 967–974.
Pierce L, Toffel MW (2013) The role of organizational scope and governance in strengthening private monitoring. *Organization Science* 24(5): 1558–1584.
Rajan B, Tezcan T, Seidmann A (2019) Service systems with heterogeneous customers: Investigating the effect of telemedicine on chronic care. *Management Science* 65(3): 1236–1267.
Reid DJ, Reid FJ (2005) Online focus groups: An in-depth comparison of computer-mediated and conventional focus group discussions. *International Journal of Market Research* 47(2): 131–162.
Ronen, J (2010) Corporate audits and how to fix them. *Journal of Economic Perspectives* 24(2): 189–210.
Sampson SE, Froehle CM (2006) Foundations and implications of a proposed unified services theory. *Production and Operations Management* 15(2): 329–343.
Short JL, Toffel MW, Hugill AR (2016) Monitoring global supply chains. *Strategic Management Journal* 37(9): 1878–1897.
Sun S, Lu SF, Rui H (2020) Does telemedicine reduce emergency room congestion? Evidence from New York State. *Information Systems Research* 31(3): 972–986.
Tan TF, Netessine S (2020) At your service on the table: Impact of tabletop technology on restaurant performance. *Management Science* 66(10): 4496–4515.

Taylor FW (1911) *The Principles of Scientific Management* (Harper, New York).
Tyre MJ, von Hippel E (1997) The situated nature of adaptive learning in organizations. *Organization Science* 8(1): 71–83.
US Bureau of Labor Statistics (2022) US business response to the COVID-19 pandemic—2021. Report, US Department of Labor, Washington, DC. https://www.bls.gov/news.release/pdf/covid2.pdf.
US Department of Health and Human Services (2024) Conducting remote regulatory assessments—Questions and answers. Report, Food and Drug Administration, *Federal Register* 89(18), January 26. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/conducting-remote-regulatory-assessments-questions-and-answers
US Department of Labor (2021) COVID-19: Increased worksite complaints and reduced OSHA inspections leave US workers' safety at increased risk. Report, Office of Inspector General—Office of Audit, US Department of Labor, Washington, DC. https://www.oig.dol.gov/public/reports/oa/2021/19-21-003-10-105.pdf.
Vargo SL, Lusch RF (2004) Evolving to a new dominant logic for marketing. *Journal of Marketing* 68(1): 1–17.
Xue M, Hitt LM, Chen P (2011) Determinants and outcomes of internet banking adoption. *Management Science* 57(2): 291–307.
Yang L, Holtz D, Jaffe S, Suri S, Sinha S, Weston J, Joyce C, Shah N, Sherman K, Hecht B, Teevan J (2022) The effects of remote work on collaboration among information workers. *Nature Human Behaviour* 6(1): 43–54.

Table 1. Summary Statistics

	Mean	SD	Min	Max
Violations	1.42	1.65	0	58
Remote audit	0.39	0.49	0	1
Audit team	0.20	0.40	0	1
Average prior in-person site exposure	2.89	3.22	0	48
Average prior in-person site exposure (log) [a]	1.10	0.71	0	3.89
Focal standard advanced training	0.20	0.40	0	1
Staff-days	1.56	1.07	0.50	30
Staff-days (log)	0.29	0.53	-0.69	3.40
Prior remote site exposure	0.15	0.35	0	1
Maximum auditing experience	344.17	211.02	0	1048
Maximum auditing experience (log) [a]	5.53	1.00	0	6.96
Percent outsourced	0.27	0.42	0	1
Multi-standard audit	0.10	0.30	0	1
Maximum management standard focus	0.56	0.29	0	1
Audit sequence [b]	6.52	3.63	1	16

N = 35,247 audits.

[a] We avoid missing logged values by adding 1 to the raw value before taking the natural log.

[b] Winsorized at the 95th percentile.

Table 2. Correlations

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1)	Violations	1.00									
(2)	Remote audit	-0.12	1.00								
(3)	Audit team	0.02	-0.12	1.00							
(4)	Average prior in-person site exposure (log)	-0.01	-0.34	0.06	1.00						
(5)	Focal standard advanced training	0.04	-0.08	0.65	0.02	1.00					
(6)	Staff-days (log)	0.17	-0.05	0.48	0.02	0.42	1.00				
(7)	Prior remote site exposure	-0.05	0.28	-0.05	0.13	-0.07	-0.01	1.00			
(8)	Maximum auditing experience (log)	-0.10	0.03	0.19	0.32	0.05	-0.05	0.15	1.00		
(9)	Percent outsourced	-0.02	-0.05	0.08	0.05	0.00	0.09	-0.04	-0.15	1.00	
(10)	Multi-standard audit	-0.04	-0.06	0.22	0.11	0.19	-0.07	-0.07	0.11	0.01	1.00
(11)	Maximum management standard focus	0.03	0.05	-0.05	-0.20	0.01	0.08	-0.08	-0.24	0.03	-0.22

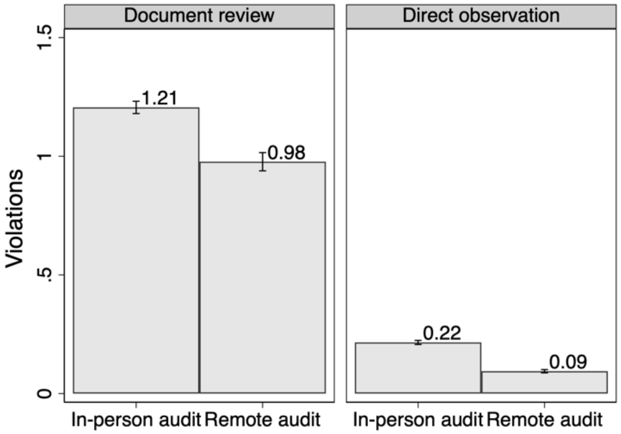
N = 35,247 audits.

Table 3. Regression Results

	Dependent variable: Violations				
	(1)	(2)	(3)	(4)	(5)
Remote audit	-0.294**	-0.210**	-0.244**	-0.393**	-0.381**
	(0.027)	(0.031)	(0.028)	(0.040)	(0.042)
Remote audit x Direct-observation–based clauses		-0.617**			
		(0.038)			
Remote audit x Audit team			-0.299**		
			(0.046)		
Remote audit x Average prior in-person site exposure (log)				0.087**	
				(0.026)	
Remote audit x Maximum management standard focus					0.148**
					(0.055)
Direct-observation–based clauses		-1.722**			
		(0.020)			
Audit team	0.031	0.051	0.131**	0.026	0.024
	(0.040)	(0.047)	(0.044)	(0.040)	(0.041)
Average prior in-person site exposure (log)	-0.141**	-0.135**	-0.142**	-0.201**	-0.142**
	(0.021)	(0.023)	(0.021)	(0.028)	(0.021)
Maximum management standard focus					0.065
					(0.059)
Focal standard advanced training	0.041	0.031	0.043	0.042	0.035
	(0.034)	(0.041)	(0.034)	(0.034)	(0.034)
Staff-days (log)	0.421**	0.406**	0.422**	0.422**	0.420**
	(0.051)	(0.066)	(0.052)	(0.051)	(0.051)
Prior remote site exposure	-0.023	-0.025	-0.016	-0.037	-0.015
	(0.029)	(0.032)	(0.029)	(0.029)	(0.029)
Maximum auditing experience (log)	-0.030*	-0.037**	-0.030*	-0.026*	-0.026*
	(0.012)	(0.013)	(0.012)	(0.012)	(0.012)
Percent outsourced	-0.177**	-0.203**	-0.175**	-0.181**	-0.180**
	(0.030)	(0.033)	(0.030)	(0.030)	(0.030)
Multi-standard audit	-0.022	-0.030	-0.011	-0.027	-0.021
	(0.042)	(0.050)	(0.042)	(0.042)	(0.041)
Site-standard fixed effects	Yes	Yes	Yes	Yes	Yes
Audit-sequence dummies	Yes	Yes	Yes	Yes	Yes
Month-year dummies	Yes	Yes	Yes	Yes	Yes
Observations (audits)	35,247		35,247	35,247	35,247
Observations (detection-mode–audits)		63,908			
Number of site–standard dyads	16,986	15,388	16,986	16,986	16,986
Number of sites	14,615	13,296	14,615	14,615	14,615

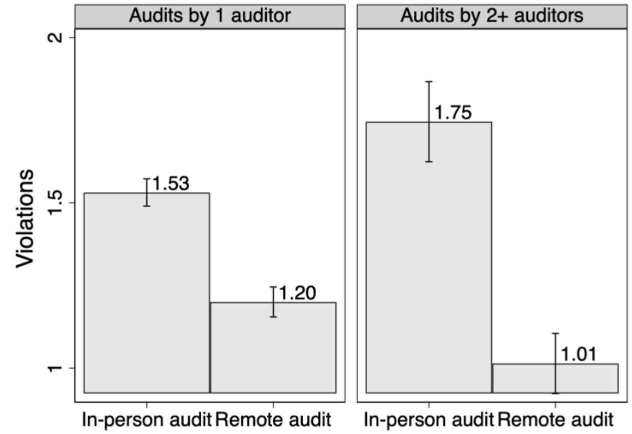
Poisson regression coefficients with standard errors are clustered by audited site; **p<0.01; *p<0.05; +p<0.10. The unit of analysis for Models 1 and 3–5 is an audit on a given date of a standard at a site and for Model 2 is an audit on a given date of a standard at a site using one of two detection-modes (direct observation or document review, the omitted category). Model 2's observation count is less than twice that of the other models because it excludes ISO 13485 audits (for which we were unable to obtain data to map its clauses to primary violation-detection mode)

Figure 1. Average Predicted Violations for Remote vs. In-person Audits, by Violation-detection Mode



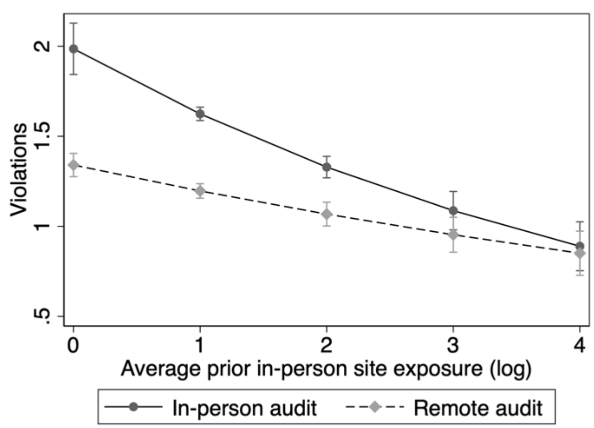
Predictive average effects and their 95% confidence intervals, based on the results of Model 2 in Table 3. The decrement of remote audits for violations identified primarily through direct observation of site processes is significantly larger than the decrement of remote audits for violations identified primarily through document review.

Figure 2. Average Predicted Violations for Remote vs. In-person Audits, by Single Auditors versus Auditor Teams



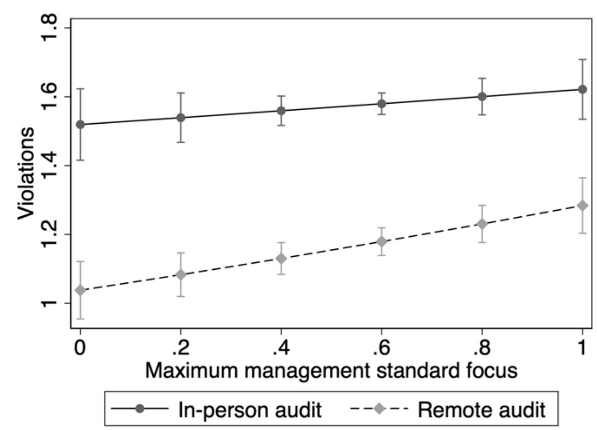
Predictive average effects and their 95% confidence intervals, based on the results of Model 3 in Table 3. The gap in predicted average violations between remote and in-person audits is significantly larger for auditor teams than for single auditors.

Figure 3. Average Predicted Effects of Remote vs. In-person Audits, by Level of Prior In-person Site Exposure



Predictive average effects and their 95% confidence intervals, based on the results of Model 4 in Table 3. The audit quality gap been remote and in-person auditing is attenuated when audit teams have greater prior in-person site exposure. Note that the x-axis depicts a log scale to match the logged variable *average prior in-person site exposure (log)*.

Figure 4. Average Predicted Violations for Remote vs. In-person Audits, by Level of Management Standard Focus



Predictive average effects and their 95% confidence intervals, based on the results of Model 5 in Table 3. The audit-quality gap been remote and in-person auditing is attenuated when audit teams have had greater focus on the management standard being audited.

Online Appendix

Table A-1. Sample Construction

Sample refinement step	Description	# audits dropped in this step	# audits remaining after this step
1. Raw data	All surveillance audits of the focal six standards in 2019–2021.		125,727
2. Purging data anomalies	Omit audits with data anomalies that resulted from data entry errors or did not reflect the typical course of audits according to the auditing company that provided our data.	– 6,657	119,070
3. Purging less-typical auditing practices	Omit sites coded as having two or more surveillance audits in a given year because, according to the auditing company that provided our data, such audits tend to have a smaller scope than annual surveillance audits.	– 28,420	90,650
4. Dropped by fixed-effects Poisson regression	Our primary Poisson regression specification with site-standard fixed effects drops observations that (a) are singleton site–standard dyads (that is, an establishment had only one surveillance audit for a particular standard during our sample period) or (b) pertain to site–standard panels whose every audit reported zero violations. (These 55,403 audits are included in the estimation sample of our robustness test specifications that omit site-standard fixed effects or replace them with random effects.)	– 55,403	35,247
Final estimation sample			35,247

.

Table A-2. Identification Test: "Cleaner" Sites are Not More Likely to Be Assigned to Remote Audits

	Dependent variable: Remote audit	
	(1) Logit	(2) OLS
Violations (prior audit)	0.058+	0.005
	(0.030)	(0.005)
Audit team (prior audit)	0.303	0.014
	(0.217)	(0.032)
Average prior in-person site exposure (log) (prior audit)	0.016	-0.027+
	(0.102)	(0.014)
Focal standard advanced training (prior audit)	-0.109	-0.010
	(0.165)	(0.024)
Staff-days (log) (prior audit)	-0.054	0.013
	(0.178)	(0.027)
Prior remote site exposure (prior audit)	-0.937**	-0.057*
	(0.225)	(0.027)
Maximum auditing experience (log) (prior audit)	-0.013	0.012
	(0.074)	(0.010)
Percent outsourced (prior audit)	0.218	0.031
	(0.164)	(0.025)
Multi-standard audit (prior audit)	-0.726**	-0.115**
	(0.150)	(0.021)
Site-standard fixed effects	Yes	Yes
Audit-sequence fixed effects	Yes	Yes
Month-year fixed effects	Yes	Yes
Observations (audits)	4,741	21,687

Logistic (Model 1) and OLS (Model 2) regression coefficients with standard errors clustered by audited site in parentheses; **p<0.01; *p<0.05; +p<0.10. The unit of analysis is an audit of a site for a particular standard.

Table A-3. Robustness Test: Controlling for Audit Team's Minimum Experience

	Dependent variable: Violations				
	(1)	(2)	(3)	(4)	(5)
Remote audit	-0.299**	-0.215**	-0.248**	-0.400**	-0.386**
	(0.027)	(0.031)	(0.028)	(0.040)	(0.042)
Remote audit x Direct-observation–based clauses		-0.617**			
		(0.038)			
Remote audit x Audit team			-0.300**		
			(0.046)		
Remote audit x Average prior in-person site exposure (log)				0.090**	
				(0.026)	
Remote audit x Maximum management standard focus					0.147**
					(0.055)
Direct-observation–based clauses		-1.722**			
		(0.020)			
Audit team	0.007	0.023	0.107*	0.006	0.005
	(0.041)	(0.048)	(0.044)	(0.041)	(0.041)
Average prior in-person site exposure (log)	-0.150**	-0.145**	-0.150**	-0.211**	-0.151**
	(0.021)	(0.023)	(0.021)	(0.028)	(0.021)
Maximum management standard focus					0.081
					(0.059)
Focal standard advanced training	0.043	0.032	0.044	0.043	0.035
	(0.034)	(0.041)	(0.034)	(0.034)	(0.034)
Staff-days (log)	0.418**	0.403**	0.420**	0.420**	0.418**
	(0.051)	(0.066)	(0.052)	(0.051)	(0.051)
Prior remote site exposure	-0.025	-0.027	-0.018	-0.039	-0.017
	(0.029)	(0.032)	(0.029)	(0.029)	(0.029)
Minimum auditing experience (log)	-0.013	-0.016	-0.013	-0.009	-0.008
	(0.010)	(0.011)	(0.010)	(0.010)	(0.010)
Percent outsourced	-0.173**	-0.198**	-0.171**	-0.177**	-0.176**
	(0.030)	(0.033)	(0.030)	(0.030)	(0.030)
Multi-standard audit	-0.023	-0.031	-0.012	-0.028	-0.022
	(0.042)	(0.050)	(0.042)	(0.042)	(0.041)
Site-standard fixed effects	Yes	Yes	Yes	Yes	Yes
Audit-sequence fixed effects	Yes	Yes	Yes	Yes	Yes
Month-year fixed effects	Yes	Yes	Yes	Yes	Yes
Observations (audits)	35,247		35,247	35,247	35,247
Observations (detection-mode–audits)		63,908			

Poisson regression coefficients with standard errors clustered by site; **p<0.01; *p<0.05; +p<0.10. See notes for Table 3. These models differ from our primary specifications in Table 3 by controlling for the audit team's *minimum auditing experience (log)* instead of its *maximum auditing experience (log)*.

Table A-4. Robustness Test: Controlling for Month-Year x Country Fixed Effects

	Dependent variable: Violations				
	(1)	(2)	(3)	(4)	(5)
Remote audit	-0.281**	-0.216**	-0.245**	-0.395**	-0.345**
	(0.033)	(0.037)	(0.034)	(0.045)	(0.047)
Remote audit x Direct-observation–based clauses		-0.617**			
		(0.038)			
Remote audit x Audit team			-0.219**		
			(0.052)		
Remote audit x Average prior in-person site exposure (log)				0.101**	
				(0.027)	
Remote audit x Maximum management standard focus					0.108+
					(0.057)
Direct-observation–based clauses		-1.722**			
		(0.020)			
Audit team	0.093*	0.098*	0.166**	0.087*	0.083*
	(0.041)	(0.046)	(0.044)	(0.041)	(0.041)
Average prior in-person site exposure (log)	-0.150**	-0.150**	-0.150**	-0.218**	-0.151**
	(0.021)	(0.023)	(0.021)	(0.028)	(0.021)
Maximum management standard focus					0.110+
					(0.059)
Focal standard advanced training	0.035	0.038	0.033	0.036	0.027
	(0.035)	(0.040)	(0.035)	(0.035)	(0.035)
Staff-days (log)	0.452**	0.459**	0.450**	0.453**	0.449**
	(0.050)	(0.057)	(0.051)	(0.050)	(0.051)
Prior remote site exposure	-0.065*	-0.076*	-0.058+	-0.079**	-0.059+
	(0.030)	(0.033)	(0.030)	(0.030)	(0.030)
Maximum auditing experience (log)	-0.019	-0.021	-0.020	-0.016	-0.014
	(0.012)	(0.014)	(0.012)	(0.012)	(0.013)
Percent outsourced	-0.209**	-0.239**	-0.207**	-0.214**	-0.214**
	(0.030)	(0.033)	(0.030)	(0.030)	(0.030)
Multi-standard audit	-0.003	-0.003	0.004	-0.010	-0.002
	(0.041)	(0.046)	(0.042)	(0.042)	(0.041)
Site-standard fixed effects	Yes	Yes	Yes	Yes	Yes
Audit-sequence dummies	Yes	Yes	Yes	Yes	Yes
Month-year x country dummies	Yes	Yes	Yes	Yes	Yes
Observations (audits)	34,579		34,579	34,579	34,579
Observations (detection-mode–audits)		62,868			

Poisson regression coefficients with standard errors clustered by site; **p<0.01; *p<0.05; +p<0.10. See notes for Table 3. These models differ from our primary specifications in Table 3 by including the full set of interactions between month-year dummies and country dummies.

Table A-5. Robustness Test: Including Site–Standard–Detection-mode Fixed Effects in Model 2

	Dependent variable: Violations
	(1)
Remote audit	-0.233**
	(0.031)
Remote audit x Direct-observation–based clauses	-0.408**
	(0.045)
Direct-observation–based clauses	(Absorbed)
Audit team	0.053
	(0.047)
Average prior in-person site exposure (log)	-0.133**
	(0.023)
Focal standard advanced training	0.032
	(0.041)
Staff-days (log)	0.403**
	(0.066)
Prior remote site exposure	-0.032
	(0.032)
Maximum auditing experience (log)	-0.038**
	(0.013)
Percent outsourced	-0.203**
	(0.033)
Multi-standard audit	-0.028
	(0.050)
Site-standard fixed effects	(Absorbed)
Site–standard–detection-mode fixed effects	Yes
Audit-sequence fixed effects	Yes
Month-year fixed effects	Yes
Observations (detection-mode–audits)	37,631
Number of site–standard dyads	15,388
Number of sites	13,296

Poisson regression coefficients with standard errors clustered by site; **p<0.01; *p<0.05; +p<0.10. See notes for Table 3. This model differs from our primary Model 2 specification in Table 3 by including site–standard–detection-mode fixed effects rather than just site-standard fixed effects, which provides within–detection-mode estimates but results in the model mechanically dropping observations that report zero violations for all instances of a site's various audits of a given standard using a particular detection mode (e.g., if Site A's ISO 9001 audits of observation-based clauses always yielded zero violations). This model's observation count is therefore less than that of Model 2 in Table 3. In this model, the main effect of *observation-based clauses* is absorbed by the inclusion of site–standard–detection-mode fixed effects.

Table A-6. Robustness Test: Assessing Influence of Audit Team Gender Composition

	Dependent variable: Violations										Comparison (see notes)
	All models estimated on the subsample for which audit team gender composition is known										
	Primary model specifications					Primary model specifications but also control for female on audit team					
	(1a)	(2a)	(3a)	(4a)	(5a)	(1b)	(2b)	(3b)	(4b)	(5b)	
Remote audit	-0.265**	-0.191**	-0.232**	-0.384**	-0.390**	-0.267**	-0.191**	-0.234**	-0.384**	-0.392**	1.50
	(0.040)	(0.044)	(0.040)	(0.059)	(0.059)	(0.040)	(0.044)	(0.040)	(0.059)	(0.059)	[0.22]
Remote audit x Direct-observation–based clauses		-0.647**					-0.647**				n/a
		(0.048)					(0.048)				[n/a]
Remote audit x Audit team			-0.341**					-0.342**			0.02
			(0.083)					(0.083)			[0.89]
Remote audit x Average prior in-person site exposure (log)				0.095**					0.094**		1.34
				(0.035)					(0.035)		[0.25]
Remote audit x Maximum management standard focus					0.219**					0.219**	0.02
					(0.074)					(0.074)	[0.89]
Direct-observation–based clauses		-1.724**					-1.724**				
		(0.026)					(0.026)				
Audit team	0.039	0.080	0.160*	0.034	0.035	0.018	0.069	0.139+	0.014	0.017	
	(0.067)	(0.076)	(0.072)	(0.067)	(0.067)	(0.068)	(0.077)	(0.073)	(0.068)	(0.068)	
Average prior in-person site exposure (log)	-0.082**	-0.084*	-0.084**	-0.153**	-0.083**	-0.085**	-0.086*	-0.088**	-0.154**	-0.086**	
	(0.030)	(0.033)	(0.030)	(0.040)	(0.030)	(0.030)	(0.034)	(0.030)	(0.040)	(0.030)	
Maximum management standard focus					0.120					0.106	
					(0.092)					(0.092)	
Focal standard advanced training	0.092+	0.042	0.093+	0.092+	0.075	0.091+	0.041	0.091+	0.090+	0.075	
	(0.054)	(0.061)	(0.054)	(0.054)	(0.054)	(0.054)	(0.062)	(0.054)	(0.054)	(0.054)	
Staff-days (log)	0.454**	0.466**	0.461**	0.454**	0.456**	0.456**	0.466**	0.463**	0.455**	0.457**	
	(0.071)	(0.085)	(0.072)	(0.071)	(0.071)	(0.071)	(0.085)	(0.072)	(0.071)	(0.071)	
Prior remote site exposure	-0.038	-0.059	-0.036	-0.052	-0.027	-0.038	-0.060	-0.036	-0.052	-0.027	
	(0.039)	(0.043)	(0.039)	(0.039)	(0.039)	(0.039)	(0.043)	(0.039)	(0.039)	(0.039)	
Maximum auditing experience (log)	-0.037+	-0.046*	-0.040*	-0.030	-0.033	-0.034+	-0.043+	-0.037+	-0.027	-0.032	
	(0.020)	(0.023)	(0.020)	(0.020)	(0.021)	(0.020)	(0.023)	(0.020)	(0.020)	(0.021)	
Percent outsourced	0.017	-0.057	0.041	0.020	-0.011	-0.039	-0.086	-0.015	-0.035	-0.061	
	(0.225)	(0.237)	(0.237)	(0.224)	(0.226)	(0.224)	(0.237)	(0.235)	(0.223)	(0.225)	
Multi-standard audit	-0.019	-0.012	-0.010	-0.026	-0.017	-0.019	-0.013	-0.010	-0.026	-0.018	
	(0.064)	(0.070)	(0.065)	(0.064)	(0.064)	(0.064)	(0.070)	(0.064)	(0.064)	(0.063)	
Female on audit team						0.094*	0.047	0.095*	0.092*	0.087*	
						(0.043)	(0.047)	(0.043)	(0.043)	(0.044)	
Site-standard fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Audit-sequence dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Month-year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Observations (audits)	17,856		17,856	17,856	17,856	17,856		17,856	17,856	17,856	
Observations (detection-mode–audits)		37,456					37,456				

Poisson regression coefficients with standard errors clustered by site; **p<0.01; *p<0.05; +p<0.10. See notes for Table 3. The comparison column displays the Wald chi-squared values, with p-values in brackets, from tests of whether the two coefficients in the row statistically differ; in the *remote audit* row, the coefficients from Models 1a and 1b are compared; n/a indicates that coefficients on *Remote audit x Direct-observation–based clauses* in Models 2a and 2b are identical, which precludes a Wald test.

Table A-7. Robustness Test: Alternative Specifications to Expand Estimation Sample

	Dependent variable: Violations									
	Random-effects Poisson regression models					Pooled Poisson regression models				
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
Remote audit	-0.222**	-0.180**	-0.179**	-0.259**	-0.344**	-0.173**	-0.145**	-0.131**	-0.180**	-0.313**
	(0.020)	(0.022)	(0.021)	(0.028)	(0.033)	(0.020)	(0.022)	(0.021)	(0.028)	(0.032)
Remote audit x Direct-observation–based clauses		-0.562**					-0.562**			
		(0.033)					(0.033)			
Remote audit x Audit team			-0.246**					-0.232**		
			(0.037)					(0.039)		
Remote audit x Average prior in-person site exposure (log)				0.036+					0.008	
				(0.020)					(0.020)	
Remote audit x Maximum management standard focus					0.215**					0.249**
					(0.045)					(0.043)
Direct-observation–based clauses		-1.726**					-1.726**			
		(0.017)					(0.017)			
Audit team	-0.137**	-0.057*	-0.066**	-0.138**	-0.134**	-0.171**	-0.073**	-0.112**	-0.171**	-0.170**
	(0.023)	(0.025)	(0.025)	(0.023)	(0.023)	(0.023)	(0.024)	(0.025)	(0.023)	(0.023)
Average prior in-person site exposure (log)	-0.166**	-0.181**	-0.166**	-0.186**	-0.169**	-0.156**	-0.180**	-0.156**	-0.160**	-0.161**
	(0.013)	(0.013)	(0.013)	(0.017)	(0.013)	(0.012)	(0.013)	(0.012)	(0.017)	(0.012)
Maximum management standard focus					-0.156**					-0.240**
					(0.033)					(0.029)
Focal standard advanced training	-0.201**	-0.363**	-0.204**	-0.200**	-0.202**	-0.273**	-0.469**	-0.275**	-0.272**	-0.268**
	(0.022)	(0.027)	(0.022)	(0.022)	(0.022)	(0.020)	(0.023)	(0.020)	(0.020)	(0.020)
Staff-days (log)	0.661**	0.645**	0.661**	0.661**	0.664**	0.672**	0.660**	0.673**	0.672**	0.678**
	(0.017)	(0.018)	(0.017)	(0.017)	(0.017)	(0.017)	(0.019)	(0.017)	(0.017)	(0.017)
Prior remote site exposure	0.026	0.013	0.032	0.021	0.037+	0.050*	0.033	0.056**	0.048*	0.058**
	(0.022)	(0.023)	(0.022)	(0.023)	(0.022)	(0.021)	(0.022)	(0.021)	(0.021)	(0.021)
Maximum auditing experience (log)	-0.060**	-0.025**	-0.060**	-0.059**	-0.065**	-0.078**	-0.024**	-0.078**	-0.078**	-0.086**
	(0.007)	(0.008)	(0.007)	(0.007)	(0.008)	(0.007)	(0.008)	(0.007)	(0.007)	(0.007)
Percent outsourced	-0.257**	-0.233**	-0.255**	-0.257**	-0.255**	-0.282**	-0.236**	-0.280**	-0.282**	-0.278**
	(0.017)	(0.018)	(0.017)	(0.017)	(0.017)	(0.017)	(0.018)	(0.017)	(0.017)	(0.017)
Multi-standard audit	-0.033	-0.020	-0.032	-0.033	-0.046+	-0.050*	-0.034	-0.050*	-0.050*	-0.078**
	(0.023)	(0.024)	(0.024)	(0.024)	(0.024)	(0.025)	(0.026)	(0.025)	(0.025)	(0.025)
Site-standard random effects	Yes	Yes	Yes	Yes	Yes					
Audit-sequence dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month-year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations (audits)	90,650		90,650	90,650	90,650	90,650		90,650	90,650	90,650
Observations (detection-mode–audits)		175,258					175,258			
Number of site–standard dyads	58,575		58,575	58,575	58,575	58,575		58,575	58,575	58,575
Number of site–standard–detection-mode dyads		56,521					56,521			

Poisson regression coefficients with standard errors clustered by site; **p<0.01; *p<0.05; +p<0.10. See notes for Table 3. These models differ from our primary specifications in Table 3 by (a) replacing site-standard fixed effects with random effects or (b) omitting those fixed effects.

Table A-8. Robustness Test: Predicting Proportional Dependent Variables

Dependent variable (see notes)	(1) Violations as proportion of standard's maximum	(2) Violations as proportion of standard–detection-mode's maximum	(3) Violations as proportion of standard's maximum	(4) Violations as proportion of standard's maximum	(5) Violations as proportion of standard's maximum
Remote audit	-0.302**	-0.205**	-0.257**	-0.399**	-0.386**
	(0.029)	(0.033)	(0.030)	(0.044)	(0.046)
Remote audit x Direct-observation–based clauses		-0.585**			
		(0.041)			
Remote audit x Audit team			-0.269**		
			(0.051)		
Remote audit x Average prior in-person site exposure (log)				0.086**	
				(0.028)	
Remote audit x Maximum management standard focus					0.145*
					(0.061)
Direct-observation–based clauses		-0.447**			
		(0.021)			
Audit team	0.018	0.030	0.111*	0.013	0.010
	(0.045)	(0.057)	(0.049)	(0.045)	(0.045)
Average prior in-person site exposure (log)	-0.149**	-0.152**	-0.150**	-0.208**	-0.150**
	(0.022)	(0.025)	(0.022)	(0.030)	(0.022)
Maximum management standard focus					0.080
					(0.062)
Focal standard advanced training	0.029	0.006	0.030	0.031	0.022
	(0.037)	(0.048)	(0.038)	(0.037)	(0.038)
Staff-days (log)	0.434**	0.477**	0.432**	0.435**	0.433**
	(0.054)	(0.078)	(0.055)	(0.054)	(0.054)
Prior remote site exposure	-0.025	0.019	-0.020	-0.039	-0.017
	(0.032)	(0.036)	(0.032)	(0.032)	(0.032)
Maximum auditing experience (log)	-0.030*	-0.042**	-0.030*	-0.027*	-0.026+
	(0.013)	(0.015)	(0.013)	(0.013)	(0.013)
Percent outsourced	-0.148**	-0.188**	-0.147**	-0.152**	-0.150**
	(0.032)	(0.037)	(0.032)	(0.032)	(0.033)
Multi-standard audit	-0.050	-0.001	-0.042	-0.053	-0.050
	(0.047)	(0.059)	(0.047)	(0.046)	(0.046)
Site-standard fixed effects	Yes	Yes	Yes	Yes	Yes
Audit-sequence fixed effects	Yes	Yes	Yes	Yes	Yes
Month-year fixed effects	Yes	Yes	Yes	Yes	Yes
Observations (audits)	35,247		35,247	35,247	35,247
Observations (detection-mode–audits)		58,056			

Poisson regression coefficients with standard errors clustered by site; **p<0.01; *p<0.05; +p<0.10. See notes for Table 3. These models differ from our primary specifications in Table 3 by replacing *violations* (a count) as the dependent variable with proportional dependent variables. The dependent variable for Models 1 and 3–5 is *violations as proportion of standard's sample maximum number of violations*, calculated by dividing *violations* by the sample maximum number of violations for this standard. The dependent variable for Model 2 is *violations as proportion of standard–detection-mode's maximum,* calculated by dividing *violations* by the sample maximum number of violations for this standard–detection-mode.

Table A-9. Robustness Test: Predicting Minor Violations

	Dependent variable: Minor violations				
	(1)	(2)	(3)	(4)	(5)
Remote audit	-0.296**	-0.008	-0.246**	-0.396**	-0.373**
	(0.027)	(0.030)	(0.028)	(0.040)	(0.042)
Remote audit x Direct-observation–based clauses		-0.624**			
		(0.039)			
Remote audit x Audit team			-0.294**		
			(0.046)		
Remote audit x Average prior in-person site exposure (log)				0.088**	
				(0.026)	
Remote audit x Maximum management standard focus					0.131*
					(0.054)
Direct-observation–based clauses		-1.706**			
		(0.020)			
Audit team	0.028	0.044	0.126**	0.022	0.020
	(0.040)	(0.049)	(0.044)	(0.040)	(0.041)
Average prior in-person site exposure (log)	-0.135**	-0.039+	-0.136**	-0.196**	-0.136**
	(0.020)	(0.022)	(0.020)	(0.027)	(0.020)
Maximum management standard focus					0.081
					(0.059)
Focal standard advanced training	0.040	0.033	0.042	0.041	0.034
	(0.034)	(0.041)	(0.035)	(0.034)	(0.035)
Staff-days (log)	0.431**	0.220**	0.433**	0.432**	0.429**
	(0.052)	(0.075)	(0.053)	(0.051)	(0.052)
Prior remote site exposure	-0.012	-0.006	-0.005	-0.026	-0.004
	(0.029)	(0.031)	(0.029)	(0.029)	(0.029)
Maximum auditing experience (log)	-0.029*	-0.019	-0.029*	-0.026*	-0.025*
	(0.012)	(0.014)	(0.012)	(0.012)	(0.012)
Percent outsourced	-0.185**	-0.079*	-0.183**	-0.189**	-0.188**
	(0.030)	(0.034)	(0.030)	(0.030)	(0.030)
Multi-standard audit	-0.018	-0.068	-0.006	-0.023	-0.017
	(0.041)	(0.055)	(0.041)	(0.041)	(0.041)
Site-standard fixed effects	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
Audit-sequence fixed effects	Yes	Yes	Yes	Yes	Yes
Observations (audits)	35,172		35,172	35,172	35,172
Observations (detection-mode–audits)		44,226			
Number of site–standard dyads	16,949	15,345	16,949	16,949	16,949
Number of sites	14,586	13,263	14,586	14,586	14,586

Poisson regression coefficients with standard errors clustered by site; **p<0.01; *p<0.05; +p<0.10. See notes for Table 3. These models differ from our primary specifications in Table 3 by replacing the dependent variable (*violations*) with *minor violations*.

Table A-10. Robustness Test: Predicting Violations of Mandatory Clauses

	Dependent variable: Violations of mandatory clauses				
	(1)	(2)	(3)	(4)	(5)
Remote audit	-0.352**	-0.227**	-0.283**	-0.440**	-0.362**
	(0.037)	(0.038)	(0.038)	(0.055)	(0.058)
Remote audit x Direct-observation–based clauses		-0.705**			
		(0.045)			
Remote audit x Audit team			-0.416**		
			(0.063)		
Remote audit x Average prior in-person site exposure (log)				0.077*	
				(0.038)	
Remote audit x Maximum management standard focus					0.017
					(0.077)
Direct-observation–based clauses		-1.247**			
		(0.023)			
Audit team	-0.040	-0.036	0.093	-0.042	-0.049
	(0.058)	(0.059)	(0.063)	(0.058)	(0.059)
Average prior in-person site exposure (log)	-0.110**	-0.108**	-0.111**	-0.163**	-0.111**
	(0.028)	(0.028)	(0.028)	(0.039)	(0.028)
Maximum management standard focus					0.099
					(0.075)
Focal standard advanced training	0.068	0.058	0.072	0.066	0.064
	(0.053)	(0.053)	(0.053)	(0.053)	(0.053)
Staff-days (log)	0.531**	0.543**	0.530**	0.529**	0.529**
	(0.079)	(0.082)	(0.080)	(0.079)	(0.079)
Prior remote site exposure	0.012	0.009	0.018	0.001	0.012
	(0.040)	(0.040)	(0.040)	(0.040)	(0.040)
Maximum auditing experience (log)	-0.036*	-0.041*	-0.035*	-0.032*	-0.033*
	(0.016)	(0.016)	(0.016)	(0.016)	(0.017)
Percent outsourced	-0.186**	-0.194**	-0.180**	-0.191**	-0.191**
	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)
Multi-standard audit	-0.023	-0.031	-0.010	-0.026	-0.022
	(0.069)	(0.070)	(0.069)	(0.069)	(0.069)
Site-standard fixed effects	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
Audit-sequence fixed effects	Yes	Yes	Yes	Yes	Yes
Observations (audits)	24,581		24,581	24,581	24,581
Observations (detection-mode–audits)		48,230			
Number of site–standard dyads	11,834	11,609	11,834	11,834	11,834
Number of sites	10,615	10,412	10,615	10,615	10,615

Poisson regression coefficients with standard errors clustered by site; **p<0.01; *p<0.05; +p<0.10. See notes for Table 3. These models differ from our primary specifications in Table 3 by replacing the dependent variable (*violations*) with *violations of mandatory clauses*.

Table A-11. Robustness Test: Subsample Excluding Sites in the European Union

	Dependent variable: Violations				
	(1)	(2)	(3)	(4)	(5)
Remote audit	-0.304**	-0.219**	-0.251**	-0.387**	-0.382**
	(0.028)	(0.032)	(0.029)	(0.042)	(0.044)
Remote audit x Direct-observation–based clauses		-0.633**			
		(0.040)			
Remote audit x Audit team			-0.298**		
			(0.047)		
Remote audit x Average prior in-person site exposure (log)				0.075**	
				(0.027)	
Remote audit x Maximum management standard focus					0.133*
					(0.057)
Direct-observation–based clauses		-1.699**			
		(0.021)			
Audit team	0.033	0.067	0.130**	0.029	0.023
	(0.042)	(0.049)	(0.046)	(0.042)	(0.043)
Average prior in-person site exposure (log)	-0.131**	-0.128**	-0.131**	-0.182**	-0.133**
	(0.022)	(0.024)	(0.022)	(0.029)	(0.022)
Maximum management standard focus					0.081
					(0.060)
Focal standard advanced training	0.052	0.034	0.056	0.052	0.046
	(0.035)	(0.042)	(0.035)	(0.035)	(0.035)
Staff-days (log)	0.428**	0.404**	0.430**	0.428**	0.427**
	(0.054)	(0.069)	(0.055)	(0.054)	(0.054)
Prior remote site exposure	-0.002	-0.013	0.004	-0.014	0.006
	(0.030)	(0.033)	(0.030)	(0.031)	(0.030)
Maximum auditing experience (log)	-0.023+	-0.029*	-0.023+	-0.020	-0.018
	(0.012)	(0.014)	(0.012)	(0.012)	(0.013)
Percent outsourced	-0.177**	-0.198**	-0.174**	-0.181**	-0.180**
	(0.031)	(0.034)	(0.031)	(0.031)	(0.031)
Multi-standard audit	-0.039	-0.048	-0.029	-0.043	-0.038
	(0.044)	(0.053)	(0.044)	(0.043)	(0.043)
Site-standard fixed effects	Yes	Yes	Yes	Yes	Yes
Audit-sequence fixed effects	Yes	Yes	Yes	Yes	Yes
Month-year fixed effects	Yes	Yes	Yes	Yes	Yes
Observations (audits)	32,260		32,260	32,260	32,260
Observations (detection-mode–audits)		58,718			

Poisson regression coefficients with standard errors clustered by site; **p<0.01; *p<0.05; +p<0.10. See notes for Table 3. These models differ from our primary specifications in Table 3 by omitting from the sample audited sites in European Union countries.

Table A-12. Robustness Test: Subsample of Site–Standard Dyads Subjected to Both In-person and Remote Audits

	Dependent variable: Violations				
	(1)	(2)	(3)	(4)	(5)
Remote audit	-0.325**	-0.266**	-0.256**	-0.419**	-0.396**
	(0.043)	(0.048)	(0.043)	(0.056)	(0.053)
Remote audit x Direct-observation–based clauses		-0.439**			
		(0.044)			
Remote audit x Audit team			-0.348**		
			(0.050)		
Remote audit x Average prior in-person site exposure (log)				0.078**	
				(0.030)	
Remote audit x Maximum management standard focus					0.125*
					(0.058)
Direct-observation–based clauses		-1.855**			
		(0.028)			
Audit team	-0.000	0.020	0.165**	-0.006	-0.009
	(0.052)	(0.062)	(0.058)	(0.052)	(0.053)
Average prior in-person site exposure (log)	-0.120**	-0.115**	-0.122**	-0.184**	-0.120**
	(0.025)	(0.027)	(0.025)	(0.036)	(0.025)
Maximum management standard focus					0.133+
					(0.074)
Focal standard advanced training	0.039	-0.007	0.050	0.040	0.024
	(0.048)	(0.059)	(0.049)	(0.048)	(0.048)
Staff-days (log)	0.466**	0.455**	0.465**	0.468**	0.463**
	(0.065)	(0.086)	(0.065)	(0.064)	(0.064)
Prior remote site exposure	-0.053	-0.046	-0.040	-0.066+	-0.042
	(0.037)	(0.040)	(0.037)	(0.037)	(0.037)
Maximum auditing experience (log)	-0.028+	-0.035*	-0.029*	-0.024	-0.025
	(0.015)	(0.017)	(0.015)	(0.015)	(0.015)
Percent outsourced	-0.165**	-0.182**	-0.163**	-0.169**	-0.172**
	(0.037)	(0.041)	(0.037)	(0.037)	(0.037)
Multi-standard audit	-0.055	-0.056	-0.034	-0.059	-0.052
	(0.057)	(0.072)	(0.058)	(0.057)	(0.057)
Site-standard fixed effects	Yes	Yes	Yes	Yes	Yes
Audit-sequence fixed effects	Yes	Yes	Yes	Yes	Yes
Month-year fixed effects	Yes	Yes	Yes	Yes	Yes
Observations (audits)	21,554		21,554	21,554	21,554
Observations (detection-mode–audits)		39,248			
Number of site–standards	10,206	9,282	10,206	10,206	10,206
Number of sites	8,986	8,187	8,986	8,986	8,986

Poisson regression coefficients with standard errors clustered by site; **p<0.01; *p<0.05; +p<0.10. See notes for Table 3. These models differ from our primary specifications in Table 3 by restricting the sample to those site–standard dyads with at least one audit conducted remotely and at least one conducted in-person.

Table A-13. Robustness Test: Subsample Restricted to Audits whose Audit Teams include at least one Auditor who Conducted Both In-Person and Remote Audits

	(1)	(2)	(3)	(4)	(5)
Remote audit	-0.289**	-0.206**	-0.241**	-0.385**	-0.367**
	(0.027)	(0.031)	(0.028)	(0.041)	(0.043)
Remote audit x Direct-observation–based clauses		-0.611**			
		(0.039)			
Remote audit x Audit team			-0.287**		
			(0.046)		
Remote audit x Average prior in-person site exposure (log)				0.084**	
				(0.026)	
Remote audit x Maximum management standard focus					0.132*
					(0.055)
Direct-observation–based clauses		-1.728**			
		(0.021)			
Audit team	0.041	0.057	0.137**	0.036	0.031
	(0.041)	(0.048)	(0.044)	(0.041)	(0.041)
Average prior in-person site exposure (log)	-0.133**	-0.129**	-0.135**	-0.191**	-0.135**
	(0.021)	(0.023)	(0.021)	(0.029)	(0.021)
Maximum management standard focus					0.104+
					(0.059)
Focal standard advanced training	0.041	0.035	0.043	0.041	0.034
	(0.035)	(0.042)	(0.035)	(0.035)	(0.035)
Staff-days (log)	0.423**	0.410**	0.425**	0.424**	0.421**
	(0.052)	(0.067)	(0.053)	(0.052)	(0.052)
Prior remote site exposure	-0.028	-0.031	-0.021	-0.041	-0.021
	(0.030)	(0.032)	(0.030)	(0.030)	(0.030)
Maximum auditing experience (log)	-0.032**	-0.039**	-0.032**	-0.028*	-0.027*
	(0.012)	(0.014)	(0.012)	(0.012)	(0.013)
Percent outsourced	-0.191**	-0.217**	-0.189**	-0.195**	-0.195**
	(0.031)	(0.034)	(0.031)	(0.031)	(0.031)
Multi-standard audit	-0.024	-0.033	-0.013	-0.028	-0.023
	(0.042)	(0.050)	(0.042)	(0.042)	(0.042)
Site-standard fixed effects	Yes	Yes	Yes	Yes	Yes
Audit-sequence fixed effects	Yes	Yes	Yes	Yes	Yes
Month-year fixed effects	Yes	Yes	Yes	Yes	Yes
Observations (audits)	34,660		34,660	34,660	34,660
Observations (detection-mode–audits)		63,200			

Poisson regression coefficients with standard errors clustered by site; **p<0.01; *p<0.05; +p<0.10. See notes for Table 3. These models differ from our primary specifications in Table 3 by restricting the sample to those audits conducted by auditors who conducted both remote and in-person audits during our sample.

Table A-14. Robustness Test of Model 2: Adding Interview-based Detection-mode Observations

	Dependent variable: Violations
	(1)
Remote audit	-0.223**
	(0.030)
Remote audit x Direct-observation–based clauses	-0.617**
	(0.038)
Remote audit x Interview-based clauses	0.033
	(0.038)
Direct-observation–based clauses	-1.722**
	(0.020)
Interview-based clauses	-2.326**
	(0.022)
Audit team	0.055
	(0.047)
Average prior in-person site exposure (log)	-0.133**
	(0.023)
Focal standard advanced training	0.035
	(0.040)
Staff-days (log)	0.408**
	(0.066)
Prior remote site exposure	-0.025
	(0.031)
Maximum auditing experience (log)	-0.035**
	(0.013)
Percent outsourced	-0.194**
	(0.032)
Multi-standard audit	-0.037
	(0.049)
Site-standard fixed effects	Included
Year fixed effects	Included
Audit-sequence fixed effects	Included
Observations (detection-mode–audits)	98,964
Number of site–standards	15,891
Number of sites	13,641

Poisson regression coefficients with standard errors clustered by site; **p<0.01; *p<0.05; +p<0.10. See notes for Table 3. The unit of analysis is the site–standard–detection-mode. This model differs from our primary Model 2 specification in Table 3 by including for every audit a third observation where violations refers to those of the clauses primarily detected by interviews, the third violation-detection mode.