

ISSN 1936-5349 (print)
ISSN 1936-5357 (online)

HARVARD

JOHN M. OLIN CENTER FOR LAW, ECONOMICS, AND BUSINESS
FELLOWS' DISCUSSION PAPER SERIES

DOES COMPLIANCE TRAINING DECREASE CORPORATE MISCONDUCT? EVIDENCE FROM FIELD DATA

Jihwon Park

Discussion Paper No. 89

05/2020

Harvard Law School
Cambridge, MA 02138

This paper can be downloaded without charge from:

The Harvard John M. Olin Fellow's Discussion Paper Series:
http://www.law.harvard.edu/programs/olin_center

Does Compliance Training Decrease Corporate Misconduct?
Evidence from Field Data

Jihwon Park

Harvard Business School

April 2020

Abstract

Firms spend significant resources on compliance training, but it is often criticized as being cosmetic. Using proprietary records on compliance training and allegations of misconduct from a large multinational firm, I investigate whether compliance training decreases corporate misconduct. I find that in-person training impacts employee behavior but do not find evidence for video training, which indicates that compliance training can be effective when employees are attentive. However, this effect lasts for only two months, suggesting that compliance training only temporarily raises awareness. I also find that the effectiveness of training is curtailed by employees' economic incentives to misbehave, such as high performance pressure and weak public enforcement. Overall, this study advances our understanding of how and when compliance training can impact employee behavior and describes its limitations.

1. Introduction

Corporate misconduct is costly. In 2018, 4.5% of the firms listed on the NYSE or NASDAQ (221 out of 4,406) faced securities class-action suits, for which \$330 billion of market capitalization was lost (SSCAC, 2019). This already significant cost becomes even larger when misconduct that is not uncovered or sanctioned by the regulators is considered. For example, Dyck, Morse, and Zingales (2017) estimate that the cost of undetected corporate misconduct is more than twice that of detected misconduct. Given these high costs, firms and the government direct significant attention and resources toward combatting misconduct.

In this paper, I investigate the effect of one of the most ubiquitous practices that firms employ to prevent corporate misconduct: compliance training. I define corporate misconduct as violation of rules by employees for which firms can be held criminally or civilly liable, and examine nine different types (as shown in Table 1). Understanding whether compliance training works matters because this training is costly for firms, as it entails not only direct costs but consumes significant employee time. Furthermore, from a regulatory perspective, compliance training is often considered a factor in firm sentencing decisions without a clear understanding of its actual impact on employee behavior. As a result, it is often criticized as being window dressing that allows firms to escape legal responsibility (Stansbury and Barry 2007; Laufer 2008; Chen and Soltes 2018). These criticisms call for empirical evidence on the effectiveness of compliance training.

Compliance training can decrease misconduct in two ways. First, it can prevent misconduct by providing clear information about the legal boundaries (the *knowledge effect*). Sometimes employees fail to comply because they do not know whether their actions are violations. By informing them, training can enhance their ability to discern misconduct and help them avoid it.

Second, training can prevent misconduct by raising awareness on compliance issues (the *awareness effect*). Theory on bounded ethicality suggests that individuals can behave in ways inconsistent with their beliefs because morality is constrained in systematic ways that favor self-serving perceptions (Banaji, Bazerman, and Chugh, 2003; Chugh, Bazerman, and Banaji 2005). For example, although mortgage lenders believe that they should not discriminate against borrowers, they may discriminate due to unconscious favoritism (Messick 1994). Training can prevent this practice by raising awareness which helps employees better reflect on their behavior and identify deviant conduct.¹

However, compliance training may not decrease misconduct for two reasons. One is that it requires employee attention; firms can mandate employees to take the training but cannot force them to heed it. In fact, anecdotal evidence suggests that employees often do not pay attention to or deeply engage with the training materials (Cohn 2012). Another is that training requires employees to voluntarily change their behavior, in contrast to other control mechanisms which focus on constraining their behavior. Prior research questions whether compliance can be taught; instead, it may be determined by an individual's trait which is unlikely to change (e.g., Davidson, Dey and Smith 2015) or by economic incentives (Becker 1968).

One difficulty in assessing whether compliance training is effective is data availability. Internal company training is not publicly observable, and misconduct is also not observed unless it has been detected by the regulators. To overcome this problem, I obtain a proprietary dataset from a large multinational manufacturer (hereafter "ManufacturingCo"), which contains information on every in-person and video compliance training given to its employees. It also contains data on allegations of misconduct reported through internal systems and whistleblowing

¹ However, in contrast to the *knowledge effect* which is considered to persist (e.g., Bandura 1977), the *awareness effect* can dissipate over time (e.g., Richards 1999; Tenbrunsel and Messick 2004)

hotlines, which enables me to measure misconduct more accurately, compared to prior research that relies on publicly detected misconduct (e.g., Dyck, Morse, and Zingales 2010; Bowen, Call, Rajgopal 2010; Heese and Pérez Cavazos 2019).² The sample period is five years, from 2013 to 2017, and the data includes information on every country in which ManufacturingCo operates (almost 100 countries in total). Both the training and allegations data cover a wide range of types of misconduct.

To test how compliance training affects the level of misconduct, I run regressions with panel data at the country-topic-time level. This design tests how compliance training on a topic given in a country relates to the level of misconduct on the topic in that country (e.g., a training on anti-bribery in the United States should affect the level of bribery in the United States). The dependent variable, *Misconduct*, equals the number of allegations that are substantiated by ManufacturingCo's internal investigations. The independent variable of interest measures the number of employees who receive training. As research finds that the effect of ethics training can be as short as four weeks (Richards 1999), I choose one month as the unit of time and include the independent variables over multiple months to investigate the longevity of the training effect. Finally, I include country-topic fixed effects and time fixed effects to control for other factors which determine the relation between misconduct and training, such as culture and time trend.

A concern with my analysis is that the relation between compliance training and misconduct may be endogenous; the decision to give training can be determined by the level of prior misconduct or other factors that are associated with it. However, ManufacturingCo conducts its training based on a plan set every three calendar years. To the extent that ManufacturingCo cannot predict where and when each type of misconduct will occur, this feature of my setting

² This is because most misconduct is only internally reported and is not uncovered externally (Near and Miceli 1985; Iwasaki 2018; Stubben and Welch 2018).

mitigates the endogeneity concern. In addition, I conduct placebo tests and granger causality test to support that my findings are not driven by omitted variables bias or reverse causality.

In my regression results, I find no association between compliance training and misconduct. The result is consistent across regressions with different levels of fixed effects and across various research designs. This suggests that compliance training does not decrease misconduct on average.

Accordingly, I explore whether the two aforementioned reasons that may cause compliance training to be ineffective explain my null findings. First, I analyze two training characteristics—training medium (in-person or video) and the size of in-person training class—to examine if inattention to training materials hinders training effectiveness. In-person training, especially those given in small classes, can better capture employees' attention than video training by facilitating interpersonal interaction and trust (e.g., Bandura 1977; Harrison 1992). Therefore, I predict that in-person training (especially if given in small class) can decrease misconduct if employee attention is an important determinant of compliance training effectiveness. The setting that ManufacturingCo provides identical training material regardless of the training medium, training class size, or country in which it is given enables me to test the effect of these characteristics. Moreover, understanding how much attention matters is also managerially significant because in-person training is costlier than video training.

Second, I examine if the expected net benefits from misconduct hinders training effectiveness. When employees make compliance decisions with the information and awareness obtained from training, they may sometimes find that compliance requires the sacrifice of economic benefits. In these situations, the training effect can be smaller. I examine two sources of variation: performance pressure and public enforcement strength. In the presence of performance pressure, benefits of misconduct are bigger, as employees are incentivized to achieve performance

often at the expense of violating rules (e.g., Richardson, Tuna, and Wu 2003; Ernst and Young 2016). Public enforcement strength, on the other hand, represents a cost of misconduct, because employees expect higher legal and reputational penalties from their violations.

I find that in-person training, especially those given in small classes, is associated with a lower level of misconduct, while I do not find evidence for video training. This suggests that compliance training can be effective when employee attention is addressed. However, the economic magnitude of in-person training effect is small, lasting for only two months and requiring almost 20,000 employees to be trained to prevent one case of misconduct. This implies that compliance training only temporarily raises awareness. Moreover, I find that in-person training effect is smaller when there is more performance pressure and in countries with weak public enforcement, indicating that economic incentive for misconduct is another obstacle for effective compliance training. Together, these results show that, although compliance training can help, companies that use it still face significant challenges in ensuring that their employees follow rules.

My results are robust to other research designs. Because my dependent variable is right-skewed count data, OLS, which assumes a linear relationship between the independent and dependent variables, may not be suitable. Thus, I also conduct tests using a zero-inflated negative binomial (ZINB) model, which is a nonparametric model developed to address data with excessive zeros in the dependent variable (Lambert 1992). I also try testing with logged variables and winsorized values. My results are robust to these alternative choices.

This paper makes several contributions. First, it contributes to the literature on misconduct (see Amiram et al. 2018 for a review of literature). Given the high cost of misconduct, many papers try to understand how it can be prevented. For instance, some investigate the effects and challenges of regulatory efforts, such as investigations by government agencies (Kedia and Rajgopal 2011;

Holzman, Marshall, and Schmidt, 2019; Nguyen 2018) and whistleblower provisions (Baloria, Marquardt, and Wiedman 2015; Call et al. 2018; Iwasaki 2018; Berger and Lee 2019; Wiedman and Zhu 2018), on combatting misconduct. This paper examines the role of compliance training, complementing concurrent research on internal governance mechanisms to reduce misconduct by Heese and Pérez Cavazos (2019), who analyze the effect of internal monitoring, and Stubben and Welch (2018) and Soltes (2018b), who examine internal whistleblowing systems.

Second, this study contributes to the literature on internal controls. COSO framework states that the three objectives of internal controls are effective operation, reliable financial reporting, and compliance with laws, regulations and policies (COSO 1992). However, prior research mostly focuses on its financial reporting consequences (e.g., Doyle, Ge, and McVay 2007b; Ashbaugh-Skaife et al. 2007; Ashbaugh-Skaife et al. 2008; Dhaliwal et al. 2011). This paper illuminates whether internal controls encourage broader compliance.

Third, this research has regulatory implications. Currently, the Federal Sentencing Guidelines for Organizations (FSGO) state that a court should decide to what degree a firm is responsible for misconduct based on its compliance training. This is based on the belief that private enforcement systems can be more efficient than public ones (Polinsky and Shavell 1993) and assumes that compliance can be taught by training. However, the lack of empirical analysis of the effectiveness of compliance training raises doubts about current practices. In fact, my findings show that compliance training prevents misconduct only in limited circumstances. This result calls into question whether measuring effectiveness based on firms' inputs is justifiable and suggests the need for a more scientific way of measuring effectiveness.

Lastly, my study may provide guidelines for firms on how they can improve their compliance training. My results show that some forms of training work better than others and has

more impact in environments with low performance pressure or strong public enforcement strength. These findings may help firms to better allocate their resources.

The rest of the paper is organized as follows. Section 2 describes institutional details and develops the hypotheses. Section 3 describes the setting and the data, provides descriptive statistics, and explains the research design. Section 4 presents results from the main analysis and cross-sectional tests. Section 5 explains additional tests and robustness tests results. Section 6 concludes.

2. Institutional background and hypotheses development

2.1. Institutional background

Formalized compliance programs,³ including compliance training, evolved in response to several large national scandals, such as the 1960s antitrust scandal and the 1970s bribery scandal. These scandals were accompanied by severe legal repercussions; for example, Lockheed Corp, one of the major firms involved in the 1970s bribery affair, paid over \$600,000 in fines for its wrongdoing (Babcock 1979). In addition, firms involved in these scandals faced considerable reputational costs, such as decreased sales, depressed stock prices, and increased cost of capital (Karpoff, Lee, and Martin 2008a). For Lockheed, much costlier than the fine was the loss of a \$1.3 billion order for new airplanes in Japan (Halloran 1976). As these legal and reputational costs were severe, firms began to create compliance programs to prevent misconduct. (See Haugh 2017 for a detailed history.)

Compliance programs became even more widespread in the 1990s when the legal incentive to establish them was strengthened. In the Federal Sentencing Guidelines for Organizations (FSGO), the United States Sentencing Commission (USSC) stated that sentencing can be reduced

³ Compliance programs encompass all personnel and systems designed to prevent and detect crimes (USSC 2004). This includes training as well as whistleblower hotlines, internal audits, internal investigations, monitoring, etc.

by up to over 90% (USSC 2004, §8C2.5. Culpability Score) or even eliminated for a firm facing federal criminal charges, if the firm can prove that it has an effective compliance program. Accordingly, compliance programs, while not mandated, have become powerful insurance against legal liability. This new regulation quickly made the establishment of compliance programs a priority for firms seeking to manage legal risks. Along with this change, clearer guidelines on compliance programs and internal controls provided by the Internal Control-Integrated Framework (published by the Committee of Sponsoring Organizations of the Treadway Commission (COSO)) accelerated its proliferation.

Compliance training is considered a key component of compliance programs because knowledge on compliance is necessary for compliant behavior. Haugh (2017) describes that compliance training is the first step out of the three in which compliance program operates: (1) education, (2) monitoring, and (3) enforcement. Consistent with this notion, when the FSGO was amended in 2004 to provide more detailed elements of an effective compliance program, it specified training as a major method of prevention (USSC 2004, §8B2.1. Effective Compliance and Ethics Program).

Owing to such legal and reputational incentives, firms are making significant investments in compliance training. The size of the global corporate compliance training market is estimated to be \$4.68 billion as of 2017 and is expected to grow up to \$7.79 billion by 2021 (Technavio 2018). Survey evidence also suggests that investment on compliance training is increasing (Thomson Reuters 2018). On top of this cost, thousands of hours of employee time per firm is spent every year on compliance training.

Despite the large cost, publics doubt about its efficacy and some suspect that firms use compliance training to escape legal responsibility. One reason for such skepticism is that

sentencing decisions are often made based on a firm's *intentions* by observing its *inputs* (e.g., how many employees watched training videos or the existence of post-training quizzes; Soltes 2018a; Chen and Soltes 2018), rather than on its *impact* on the *output* (i.e., the level of misconduct), due to the vague definition of "effectiveness" in the law and the difficulty in measuring the impact on misconduct.⁴ Consequently, firms themselves assess and manage their compliance training based on these naïve measures as well. In a survey by Deloitte and Compliance Week of compliance officers, half of respondents indicated that they measure compliance program effectiveness through completion rates (i.e., the proportion of employees who finished watching training videos) for compliance training (Deloitte 2016).

2.2. Main hypothesis: Compliance training and the level of corporate misconduct

There are two ways in which compliance training can decrease misconduct. First, it can do so by providing clear and specific information about what is legally permitted, with detailed examples of situations employees may face (the *knowledge effect*). Law is often complex and evolving, requiring individuals and firms to stay abreast of relevant rules (Katz and Bommarito 2014). In fact, research documents that programs that highlight rule compliance, with guidelines on rules and sanctions, result in increase in compliant behavior (Weaver and Trevino 1999; Kowaleski, Sutherland and Vetter 2019).

Second, compliance training can prevent misconduct by raising awareness on compliance issues (the *awareness effect*). Even with relevant knowledge, individuals may unconsciously engage in misconduct on account of "bounded ethicality," which refers to a systematic bias that

⁴ For example, the FSGO's seven key criteria for judging a compliance program effectiveness include "*reasonable oversight* with respect to the implementation and effectiveness of the compliance and ethics program" and "*reasonable steps* to respond appropriately to the criminal conduct" (USSC 2004, emphasis added). There is no clear standard to ascertain what constitutes effectiveness, and the definition of an effective compliance program is inconsistent across different regulations (Miller 2014).

favors self-serving perceptions and leads to behaviors that contradict one's intended ethical standards (e.g., Banaji, Bazerman, and Chugh 2003). As an example, Moore, Tetlock, Tanlu, and Bazerman (2006) provide evidence that individuals engage in misconduct because they fail to recognize their own conflicts of interest, while they notice others' conflicts of interest. In this case, compliance training can prevent misconduct by helping employees recognize their own susceptibility to unconscious biases (Bazerman and Banaji 2004).

However, there are two reasons why training may not decrease misconduct.⁵ One is that compliance training requires that employees pay attention to the training. Firms' ability to discipline employees can be limited (Polinsky and Shavell 1993);⁶ although a firm can make employees attend training, it cannot force them to pay attention or understand. When Garth Peterson, a former managing director of Morgan Stanley, was charged under the Foreign Corrupt Practices Act (FCPA) for bribing Chinese government officials in 2012, and the firm itself escaped legal responsibility because it had anti-bribery and anti-corruption trainings, he accused the government of excusing the firm arguing that no one in the company took compliance training seriously. He noted: "[A]ll you have to do is say, 'Garth Peterson's on the phone,' and they check the box that says, he's complied ... and then you either quietly hang up, or you just put your phone aside and you do your other work. That was the culture" (Cohn 2012).

Another is that it requires employees to voluntarily change their behavior. Whether compliance can be induced by training is questionable. For example, it may be determined by an individual's trait which is difficult to change (e.g., Davidson, Dey and Smith 2015). Another

⁵ I note that the two reasons that I suggest may not be comprehensive. For instance, the existence of other compliance programs may mitigate the effect of training. However, this is difficult to test with data from one firm, especially one which did not undergo significant change in these programs.

⁶ Polinsky and Shavell (1993) argue that this is why it is socially optimal to publicly punish employees as well as the firm for a corporate misconduct; firms cannot impose sufficient penalties to prevent employee wrongdoing.

possibility is that it is a function of perceived net benefits from misconduct (Becker 1968). If compliance behavior is critically driven by personalities and economic considerations, training may not be able to change the level of misconduct.

2.3. Cross-sectional predictions: Level of employee attention

I examine two training characteristics—(1) the training medium (in-person versus video) and (2) the in-person training class size—to examine whether compliance training effectiveness is hindered by employees’ inattention toward training materials. Understanding how much attention matters is also important because more dedicated training is costlier.

Training medium

I expect that the impact of in-person training on an individual employee will be bigger than that of video training if employee attention is an important determinant of compliance training effectiveness. Prior research suggests that in-person training can be more effective than other medium because it better addresses recipients’ attention towards the training materials by facilitating interpersonal interaction and trust (e.g., Bandura 1977; Harrison 1992). In contrast, the isolated environment in which video training is received allows employees to ignore the training materials. Helping employees to pay attention can be especially crucial for compliance training, because there is no direct incentive for them to do so.

Understanding whether and how much employee attention matters for compliance training also has implication for firms regarding their resource allocation, because more dedicated training is costlier. Even if the impact of in-person training on an individual is larger, whether its benefit is high enough to justify its higher cost is debatable. While in-person training enables greater engagement from the participants, video training allows learning opportunity for a wider audience at a given cost, including those who face constraints to participate in-person. Moreover, video

training tools such as post-training quizzes can partly address employee attention. Therefore, which is more economically efficient is an empirical question.

In-person training class size

In the similar context, I examine if in-person training class size determines the effectiveness of compliance training. Employee attention to the communicated content is expected to be better addressed in small classes through interaction, experiences, and supervision (Bandura and Mischel 1965; Harrison 1992; Lazear 2001). Therefore, I expect that the misconduct-decreasing effect of compliance training will be larger in small class than large class if employee attention is a critical determinant of its effectiveness.

2.4. Cross-sectional predictions: Economic incentives for misconduct

Another potential hindrance to effective compliance training is economic incentives for misconduct (Becker 1968). I examine whether training effectiveness varies with two economic incentives to engage in misconduct: (1) the degree of performance pressure and (2) public enforcement strength.

Performance pressure

The expected benefits to misconduct are likely to be larger in the presence of performance pressure, and research supports this. For instance, performance pressure, such as having to meet earnings thresholds (Richardson, Tuna, and Wu 2003) or raise capital (Beneish 1999a), results in increased misconduct related to earnings management. Similarly, Christensen et al. (2017) document a trade-off between safety and productivity in coal mines, and Heese and Pérez Cavazos (2019) find that monitoring from headquarters can add pressure to local facilities and lead to more misconduct. Survey evidence also shows that ethical practices can be compromised by pressure to perform. For example, in a survey of 2,825 executives conducted by Ernst and Young, 42% of the

respondents said they could justify unethical behavior to meet financial targets (Ernst and Young 2016). Based on these evidences, I expect training to be less effective in environments with high performance pressure.

Public enforcement strength

Lastly, I examine how the effects of training vary across countries with different levels of public enforcement. In countries with weak public enforcement, the expected legal costs of misconduct are likely to be smaller because of less cooperation by the international law enforcement agencies and more tolerance for misconduct. Reputational cost is likely to follow legal judgments of wrongdoing, so it, too, is likely to be lower in these countries. Consistent with this prediction, research finds that firms provide higher quality financial reporting in strong enforcement regimes (Daske et al. 2008). Thus, I expect training to have greater impact in countries with stronger public enforcement.

3. Data and research design

3.1. Data

My analyses are based on a proprietary dataset from ManufacturingCo, which is a U.S.-based, Fortune 500 company. It is one of the world's largest manufacturers, operating in almost 100 countries and employing nearly 100,000 employees. I obtained a complete set of data on its 1) misconduct allegations, 2) compliance training, and 3) human resource information such as country-level headcount, from 2013 to 2017.

The Setting

Compliance programs at ManufacturingCo are governed by the company's compliance team, which is an independent function from the rest of the organization.⁷ There are three main roles of the compliance team: compliance training, investigations of allegations, and risk intelligence, which refers to analysis of compliance-related data (collected from investigations, audits, and reports by third parties, etc.) to plan compliance programs and manage compliance risk. Although compliance training materials and whistleblowing hotline systems are supplied by third party compliance management system providers, the compliance team decides how to give training and handle allegations.

The objective of the compliance team is to prevent public detection of misconduct. Therefore, its incentive is to minimize compliance risk by preventing misconduct through training, fair investigation and actions against the allegations to 'resolve them before they become more costly to the firm' (Stubben and Welch 2018, p1). Accordingly, the team's performance is not measured based on the number of allegations that arise nor on their substantiation decisions.

During the sample period, there was no significant change in ManufacturingCo's compliance programs; the size of the compliance team was stable, and there were no changes in the use of the whistleblowing system or internal investigation policies.

ManufacturingCo conducts two types of compliance training: in-person and video. In-person training is provided to employees at physical locations. Video training refers to video lectures (with quizzes at the end) that employees are assigned to watch before a deadline. In-person training is more often given to employees in locations close to ManufacturingCo's regional headquarters, where employee population is concentrated. On the other hand, video training is given regardless of where employees are located. In most cases (roughly 65% in my sample period),

⁷ One exception is internal audit which is conducted by the company's audit committee. ManufacturingCo's audit committee and compliance team communicate once every quarter to discuss compliance issues.

training is conducted on a specific compliance topic, such as antitrust, bribery, or employee relations. ManufacturingCo also conducts training on general moral behavior, such as business ethics, integrity, and code of conduct to encourage a culture of ethics (hereafter “ethics training”).

ManufacturingCo provides identical training material regardless of the training medium, in-person training class size, or country in which it is given, due to two reasons. First, it is too costly to cater training materials. Training already requires translation of the materials into over 30 languages which demands significant time and money, and further refinement is impracticable. Second, ManufacturingCo aims to build a compliance culture with one standard.

Compliance training at ManufacturingCo is determined on a three-year basis. The training is planned as follows. First, ManufacturingCo’s compliance team prepares a proposal for training, based on their risk intelligence analyses. Then each department validates this proposal and confirms the plan. This sets the video training schedules for the upcoming three years. For in-person training, the local legal counsel has some discretion to adjust the plans at the beginning of each calendar year, but this discretion is rarely exercised. This process determines (1) the medium of training, (2) the topic, (3) the recipients, (4) the country, and (5) the time. Moreover, training is mandatory. To summarize, training is exogenously given during the three years once the plan is confirmed.

Figure 1 provides some examples of how training was conducted in ManufacturingCo. Each black dot represents an in-person training, and each blue diamond represents a video training. The numbers below the dots (diamonds) are the number of employees who received each training, and the numbers above the arrow indicate the months, to show when the trainings were conducted. As this figure shows, the frequency of training and the number of people who receive vary across topics, countries, and time.

Allegations

Data on allegations contains a comprehensive list of concerns that has been raised about ManufacturingCo's business practices. There are several channels through which allegations arise. Both employees and external parties can use a toll-free hotline via telephone or the web, set up by Navex Global, an ethics and compliance software platform. Alternatively, they can also send an email or a paper mail to ManufacturingCo's compliance department. Employees also can report misconduct internally. When reporting, they must provide information on the country where the violation occurred, choose the topic from a list of predetermined categories, and provide a summary of the case. The reporter can choose to remain anonymous. All information described in this process—the country and topic of allegation, the summary, anonymity, the reporting method, and the date of the report are available in the data.

From the topic of concern that the reporter chooses, I make manual adjustments in three cases. First, I reclassify cases that fall into categories with unclear names. For example, one category is named "Financial and Fiscal Compliance." This category spans from embezzlement to financial misreporting. Similarly, a category named "Legal and Regulatory Violations" relates to different types of legal issues. Thus, I exclude these categories and reclassify their cases to corresponding categories. Second, I reclassify cases in which the description does not coincide with the category. For instance, an allegation that ManufacturingCo is choosing a vendor based on relationship instead of competitive bidding is clearly a case of bid rigging, but the reporter chose to report it as "Antitrust," which is supposed to refer to ManufacturingCo's anticompetitive sales activities. I manually review each case description to reclassify these. Third, I exclude allegations that describe other entities' misconduct, not ManufacturingCo's. These are reported to warn or

inform ManufacturingCo, and they are excluded because they clearly do not relate to my research question. Appendix 1 shows examples of allegations for each topic.

Using the total number of internal allegations as a proxy for misconduct can be problematic as not all allegations are proven to be cases of misconduct. Therefore, I utilize the internal investigation outcome information that is available in the data. Once a report is submitted, ManufacturingCo opens an investigation and collects relevant information. The case is closed when ManufacturingCo reaches a conclusion: the allegations are either substantiated, unsubstantiated, or closed without conclusion due to lack of evidence. (Cases without a conclusion represent less than 10%). I define the substantiated allegations as examples of misconduct. This helps me measure the level of misconduct the most conservatively, because these cases are concluded by the company to be violation of rules.

Columns (1) to (4) in Table 1 shows the frequency of allegations by topic.⁸ Column (1) shows the total number of allegations, and column (2) shows the relative frequency of allegations. Column (3) displays the occasions of misconduct (substantiated allegations), and column (4) the number of unsubstantiated allegations. Only about 43% of the total allegations represent misconduct, consistent with Dyck, Morse, and Zingales (2010)'s prediction that not all allegations are not actual violations.

Table 1 Column (3) shows that misconduct occurs much more frequently than is publicly detected. While less than 10 instances of misconduct are publicly disclosed for ManufacturingCo in the Violation Tracker database during the sample period, there are almost 1,000 instances in my

⁸ There are three categories—embezzlement, information security, and conflicts of interest—that I do not include in the analyses, as they do not qualify as *corporate misconduct*, which is the subject of this paper. Firms are not responsible for these forms of misconduct; in fact, they are the victim. The motive for providing training on these topics is to protect firm assets from employees, not to manage the firm's legal and reputational risks. See Podgor et al. (2013) and Soltes (2018c) for more discussion of corporate misconduct and *respondeat superior*.

data. However, Employee relations/HR allegations, which can be internally sufficiently addressed, comprise more than 75% of the total. Nevertheless, the frequency of misconduct identified in this dataset is still large, even when only those that the firm is likely to be criminally charged for are included. For example, there are as many as 23 misconduct cases involving bribery, and all of these would violate the FCPA, since it does not have a materiality threshold.

To better assess the economic magnitude of misconduct, I calculate the estimated legal penalty if the misconduct in my data were to be publicly revealed, based on the historical average and median penalty amounts for each type of misconduct from the Violation Tracker database. This database includes enforcement data obtained from more than 40 different federal regulatory agencies and all divisions of the Justice Department and selected types of class-action lawsuits since 2000. It categorizes each violation based on the primary offense, and I match these categories to ManufacturingCo's classifications. Appendix 2 shows the average and median penalty amount for each category. In the extreme, if all misconduct in my data were to be sanctioned, the estimated penalty for ManufacturingCo would be more than \$1 billion (using the average penalty amount; \$590 million using the median).⁹ If reputational costs were considered, the amount could be much higher (Karpoff, Lee, and Martin 2008a). This exercise shows that misconduct occurs much more frequently than we can infer from publicly disclosed sources and suggests that the potential economic cost is much higher than often estimated, in line with arguments of Dyck, Morse, and Zingales (2017) and Soltes (2018c).¹⁰

Columns (1) to (3) in Table 2 Panel A show the frequency of allegations by region. Misconduct occurs most frequently in Latin America, representing about 54% of the total.

⁹ This contributes to the sensitive nature of the dataset used in this research.

¹⁰ The misconduct in my dataset may be less severe than the ones that are in the Violation Tracker database, with lower penalty amounts than the average or the median from the database. In that case, my estimates are too high.

Allegations in Asia, the Middle East, and Africa are about 18%; North America about 20%; and Europe 9%. Meanwhile, the relative proportion of employees for each region is 24% for Latin America; 29% for Asia, the Middle East and Africa; 15% for North America; and 32% for Europe, which is inconsistent with the proportion of allegations. This suggests that there are other major factors than the number of employees that contribute to the level of misconduct. Additionally, Columns (1) to (3) in Table 2 Panel B show that there is increasing trend in the number of allegations over time. In my regression, I include fixed effects to control for time trend and other country characteristics that are associated with the likelihood of misconduct.

Compliance Training

For in-person training, the following information is available in my data: (1) the country of training, (2) the topic, (3) the date, and (4) the number of people who attended the training. For video training, I observe (1) the country of training, (2) the topic, (3) the assignment date and the deadline, (4) identifiers of each employee assigned each training, and (5) whether each employee completed the assignment.

Columns (5) to (8) in Table 1 shows the total number of employees who received training by topics. The most remarkable point is that the topics where allegations occur most frequently and those on which trainings are given are different. While 76% of allegations arise from employee relations/HR, minimal training is given on this topic. Instead, training is frequently conducted on topics where average legal penalties (and presumably reputational costs as well) are high, such as antitrust and bribery. This suggests ManufacturingCo makes training decisions based on expected monetary costs of misconduct.

In columns (4) to (7) of Table 2, Panel A, the number of employees who received training is shown by region. The ratio of video participants is similar to the employee population across

region, indicating that video training is provided evenly across regions. On the other hand, in-person training is concentrated in Latin America, where allegations most frequently occur. This suggests ManufacturingCo considers prior misconduct statistics when planning its in-person training. Columns (4) to (7) of Table 2, Panel B, show there is no systematic trend in the frequency of training though there is variation across different years. This contrasts to the increasing trend in allegations, shown in Columns (1) to (3).

3.2. Research design

To test the effect of compliance training on misconduct, I perform the following regression analysis.

$$Misconduct_{c,i,t} = \alpha + \beta \cdot \sum_{t=-3}^0 Train_{c,i,t} + Fixed\ effects + \varepsilon, \quad (1)$$

where *Misconduct* is the number of substantiated allegations and *Train* is the number of employees who received training on a topic. Subscripts c, i, and t denote country, topic, and time (month).

As the subscripts suggest, the unit of analysis is country-topic-month. I choose this unit of analysis because compliance training should affect the level of misconduct 1) in the country where the training is provided and 2) on the topic of the training (e.g., training given to employees in the United States on anti-bribery should affect the level of bribery in the United States). I exclude ethics training in my main analysis, because it does not have a particular compliance topic. (I conduct a separate test for ethics training in Section 5.6.) I choose month as the unit of time because research has shown that the effect of ethics training can be as short as four weeks (Richards 1999). To explore the duration of training effectiveness, I include training variables from three months before the allegation to the month of allegation. Lastly, I include different levels of fixed effects to control for the potential time trend in the level of misconduct and other unobservable characteristics that cause potential correlation between training and misconduct.

In the cross-sectional test that examines the differential effect of the training medium, I split *Train* into two variables, *In-person* and *Video*. *In-person* is the number of employees who received in-person training, and *Video* is the number of employees who received video training. Similarly, for in-person training class size test, I split *In-person* into two variables, *In-person_small* and *In-person_big*. *In-person_small* is the number of employees who received in-person training in small (below or equal to 12 people, which is the median class size) class, and *In-person_big* is the number of employees who received in-person training in large (above median class size) class.

To examine the effects of performance pressure, I focus on the countries where ManufacturingCo has production facilities and measure performance pressure by comparing realized productivity to its targets. ManufacturingCo uses Global Effectiveness (GE) to manage its operational efficiency and sets GE targets for each month.¹¹ Subsample groups are formed depending on whether the GE targets were missed or met on average during the sample period in the country where production facilities are located, and I compare the training effectiveness between the two groups.

In my last cross-sectional test, I examine the relation between compliance training and misconduct under different public enforcement environments. Following prior research (e.g., Daske et al. 2008), I measure public enforcement in different countries by the Rule of Law index published by the World Bank.¹² Then I perform a subsample test based on whether the Rule of Law index is above or below the median.

¹¹ GE is calculated as net operating time divided by total hours used. Total hours used is the number of hours production lines were used. Net operating time is the number of hours production lines were used to produce goods that meet quality standards and is equal to total hours used less maintenance time, time loss due to delays, material shortages, breakdowns, defective products, etc.

¹² Available at: <https://databank.worldbank.org/data/source/worldwide-governance-indicators/>

Table 3 provides the descriptive statistics of the variables. One notable observation is that my dependent variable, *Misconduct*, is skewed to the right (which reflects that misconduct rare). In this case, other generalized linear models for count data, such as a Poisson or zero-inflated negative binomial model, may be more suitable than OLS (Wooldridge 2002). However, the downside of using nonparametric models is that they suffer from an incidental parameter problem when many fixed effects are used (Lancaster 2000; Greene 2002). Thus, I prioritize using OLS with many fixed effects as my main statistical model.¹³ However, I also run zero-inflated negative binomial regressions, which can address count data with excessive zeros in dependent variables (Lambert 1992). I choose negative binomial rather than Poisson because the standard deviation is much higher (10 times) than the mean.

Figure 2 shows the average frequency of misconduct (substantiated allegations), unsubstantiated allegations, and allegations closed without conclusions from two months before compliance training to four months afterward. Panel A includes all training, Panel B includes in-person training, and Panel C video training. There is no evident pattern in the allegations in Panel A. However, when I decompose it according to the training medium in Panels B and C, this suggests a potential decrease in the frequency of misconduct in Month[1] and Month[2] around in-person training. In the next section, I run a regression analysis to test the relation more rigorously.

4. Results

4.1. Main analysis: Compliance training and the level of corporate misconduct

¹³ In this setting, multiple fixed effects used in my analysis helps mitigate the skewness problem because they subsume the groups that do not have variation in the level of misconduct. For example, my regression result with country-topic fixed effects is similar to when I run a regression with a subsample of country-topic matches that have at least one case of misconduct, in which case the skewness is smaller.

Table 4 displays the results from the regression described in Section 3.2. In column (1), I include time fixed effects and topic fixed effects to control for the potential time trend in the level of misconduct and other unobservable topic-specific characteristics that cause potential correlation between training and misconduct. I also control for the number of employees in each country to scale the level of misconduct. In column (2), I include country fixed effects, rather than controlling for the number of employees. In column (3), I include country-topic fixed effects, instead of country fixed effects and topic fixed effects. This controls for the potential difference in the level of misconduct across different country-topic groups. In column (4), results using zero-inflated negative binomial (ZINB) model using country-topic indicators and time indicators as inflators is displayed.

The coefficients on *Train* are insignificant across all four columns, suggesting that training does not affect misconduct on average. In column (1), the coefficient on *Employees* is significantly positive, verifying that misconduct increases with the scale of operations. The R-squared does not increase substantially from column (1) to column (2) but does from column (2) (10.7%) to column (3) (45.8%). This indicates that the level of misconduct is largely determined by the interaction between the country and the type of misconduct. Because the country-topic fixed effects explain much of the misconduct frequency, I use this model for rest of my OLS analyses.

4.2. Cross-sectional analyses: Level of employee attention

Training medium

Table 5 shows the results that test the differential effects of training based on the training medium. In column (1), the coefficients on $\ln - person_{t-1}$ and $\ln - person_{t-2}$ are significantly negative, suggesting that in-person training is associated with lower level of misconduct for two months, as shown in the patterns from Figure 2. In contrast, none of the coefficients on *Video* are

significant, indicating that video training is not related to the level of misconduct. The implication is similar in column (2) using ZINB model. These results suggest that training can decrease misconduct only when employee attention is addressed, and further suggest that video training may not be cost effective due to inattention. Moreover, the short duration of the in-person training effect supports the idea that compliance training decreases misconduct by raising awareness (Richards 1999; Tenbrunsel and Messick 2004).

In terms of the economic magnitude, the coefficients suggest that training 1,000 employees is associated with 0.052 (0.028 in the month after the training and additional 0.024 in the second month after the training) less cases of misconduct over the four months, from the month that the training is offered to three months after. In other words, 19,231 ($=1/0.052$) employees should be trained to prevent one misconduct. Given that 33,031 employees received in-person training (excluding ethics training) over five years, only approximately 1.7 occasions of misconduct were prevented. Considering that there were almost 1,000 cases of misconduct over the sample period, the economic significance of the in-person training does not seem to be large.

I further translate my findings to legal benefit in dollar amounts to provide a better sense of the economic significance of the result. I multiply the number of misconduct cases that was prevented (1.7) by the weighted average (median) penalty amount for all misconduct in the Violation Tracker database over the sample period (\$4.3 million (\$0.6 million)) by the probability of misconduct detection (31% from Dyck, Morse, and Zingales 2017), and obtain the legal benefit of \$2.3 million (\$0.3 million) over the five years.¹⁴

In-person training class size

¹⁴ Probability of detection is less than 1% when I calculate it as the number of misconduct cases for ManufacturingCo in the Violation Tracker database over the number of substantiated allegations in the internal records. Using this probability will lead to the legal benefit of less than \$0.1 million.

Table 6 reports regression results that test differential effects of in-person training according to class size. In both columns (1) and (2), the coefficients on *In-person_small* are larger than those on *In-person_big*. For example, according to column (1), training 1,000 employees in small classes is associated with 2.512(=0.792+0.810+0.910) fewer cases of misconduct over the three months from training. In contrast, training 1,000 employees in large classes is associated with only 0.025 fewer cases of misconduct (In other words, about 40,000 employees must be trained to prevent one case of misconduct). This result is consistent with the argument that training that offers more interaction for participants is effective (Bandura and Mischel 1965; Harrison 1992) and supports that employee attention is a key determinant of compliance training effectiveness.

4.3. Cross-sectional analyses: Economic incentives for misconduct

Performance pressure

In Table 7, the differential effect of training according to the level of performance pressure is examined. In column (1), only the coefficient on $\ln - person_{t-1}$ is significantly negative. In contrast, in column (2), both coefficients on $\ln - person_{t-1}$ and $\ln - person_{t-2}$ are significantly negative. The economic magnitude of these coefficients is much larger in column (2); the coefficients suggest that 34,482 employees should be trained to prevent one case of misconduct in environments with greater performance pressure, while it takes only 3,236 (=1,000/(0.181+0.128)) employees when there is less pressure. The coefficients for $\ln - person_{t-2}$ are statistically different between the two columns. In addition, the ZINB result that is shown in the Internet Appendix Table A1 consistently finds that training is more effective under lower performance pressure.¹⁵

¹⁵ Although the coefficient for $\ln - person_t$ is statistically significantly more negative in column (1), in contrast to my prediction, that for $\ln - person_{t-2}$ is more negative in column (2), with a much larger economic magnitude.

In a supplementary test (result shown in Table A2 of the Internet Appendix), I use employee turnover as an alternative proxy for performance pressure, following Heese and Pérez Cavazos (2019). Countries with high turnover likely face more pressure because they are poorly performing. I measure employee turnover as the number of employees who left the firm divided by the total number of employees in a given country-year. I split the sample into two subgroups, above and below the turnover median. In column (2), all the coefficients on *In-person* are significantly negative, while only the coefficient on $In - person_{t-1}$ is significant in column (1). The implication is similar to that of Table 7.

Overall, these results indicate that, although in-person training can help employees avoid committing misconduct, the effect may be weaker for employees who stand to benefit more from misconduct.

Public enforcement strengths

Table 8 shows how the relation between compliance training and misconduct varies across countries with different levels of public enforcement. Column (1) is the regression result with a subsample that includes countries with strong public enforcement (with an above-median Rule of Law index), and column (2) is the result for countries with weak public enforcement (with a below-median Rule of Law index). The coefficients on $In - person_{t-1}$ and $In - person_{t-2}$ in column (1) are significantly negative, while those in column (2) are negative yet insignificant. The coefficients for $In - person_{t-1}$ and $In - person_{t-2}$ are statistically different at 10% levels between the two columns. The ZINB results (shown in the Internet Appendix Table A3) show that in-person training decreases next month's misconduct more significantly under a strong enforcement environment. This result shows that training is less effective when the expected penalties for misconduct are low.

5. Additional analyses

5.1. Likelihood of reporting

A concern with my analysis is that my dependent variable of interest, *Misconduct*, is a joint function of the occurrence of misconduct and the reporting of misconduct and thus depends on the likelihood of reporting as well as misconduct itself. Because of this concern, Berger and Lee (2019) use F-score and M-score to measure the predicted probability of fraud, instead of detected fraud. In addition, Warren, Gaspar and Laufer (2014) find that training increases the likelihood of misconduct reporting. Thus, I do several tests to address the reporting effect.¹⁶

First, I examine how the number of unsubstantiated allegations is related to compliance training. The number of unsubstantiated allegations can be a proxy for reporting behavior because they are reported cases that are not proven to be actual misconduct. The regression result with unsubstantiated allegations as the dependent variable is shown in column (1) of Table 9. It shows that there is no statistical association between compliance training and unsubstantiated allegations.

This result can be interpreted in two ways. On the one hand, it may suggest that training does not critically affect the likelihood of reporting. In this case, *Misconduct* is not a biased estimate of misconduct. On the other hand, it could be evidence that both the *awareness effect* and the *knowledge effect* have an impact on reporting probability. This is because the *awareness effect* is expected to increase the reporting probability of unsubstantiated cases, but the *knowledge effect* will decrease it as it helps employees distinguish misconduct. In this case, my results in Section 4 are likely to be biased against my findings, as both channels are likely to increase the reporting

¹⁶ Tests in this section, as well as many of those in other parts of Section 5, are based on the training medium test because it is the primary effect that I find.

probability of misconduct (substantiated cases). The relation between the two effects of training and reporting probability is shown in Figure 3.

The second and third strategies make use of the fact that outsiders (i.e., people who are not employees of ManufacturingCo) can also blow the whistle. Because outsiders do not receive compliance training, their likelihood of reporting should not have changed. However, I cannot tease out outsider allegations perfectly, as I do not have the identity of the reporter for most cases, either because the reporter chose to remain anonymous or the identity is not mentioned in the dataset. Instead, I exclude allegations made through the “reporting up” form which is a separate internal reporting channel that only the employees can use, in contrast to whistleblowing hotlines which are accessible to both employees and outsiders. The result is shown in column (2) and are consistent to my finding in Table 5. Lastly, in column (3), I exclude two topics (employee relations/HR and security) from the sample, because the allegations on these topics are most likely to arise from employees (as they are likely to be the victims of misconduct; Stubben and Welch 2019). The result is weaker but consistent with the original finding.

5.2. Placebo tests and reverse-causality test

Next, I perform several tests to strengthen my causality argument. First, I perform a placebo test where the dependent variable is the frequency of misconduct on other issues than the training topic. Putting this formally,

$$\sum_{j \neq i} Misconduct_{c,j,t} = \alpha + \beta \cdot \sum_{t=-3}^0 In-person_{c,i,t} + \gamma \cdot \sum_{t=-3}^0 Video_{c,i,t} + Fixed\ effects + \varepsilon, \quad (2)$$

The intuition is that, if the relationship between training and misconduct is causal, training should relate only to the level of misconduct on the training topic training and not to other types of

misconduct (e.g., training on bribery should affect only the level of bribery). The result is shown in Table 10 column (1). All coefficients are insignificant, confirming my conjecture.

In column (2), I conduct another placebo test, where the dependent variable is $Misconduct_{t+4}$, which is the frequency of misconduct on the same topic in the same country as the training but four months after the training. I choose $t+4$ because I include four months of training, from $t-3$ to t , as my independent variable assuming that training effect can last for a couple of months, and $t+4$ is the first month when I assume that the effect of training in month t will not affect misconduct. Again, all coefficients are statistically insignificant in this regression.¹⁷

In column (3), I include $In - person_{t+1}$ and $In - person_{t+2}$ to test for reverse causality. If ManufacturingCo decides to give training in a timely manner based on its observations of misconduct, I expect to see positive coefficients on these variables. The result in columns (3) show that the coefficients on these variables are insignificant, indicating that the decision to conduct training is not made based on the recent occurrences of misconduct. This is consistent with the fact that ManufacturingCo plans training in advance on a long-term basis. Overall, these results provide additional evidence that endogeneity concerns are not significant in my results.

5.3. Addressing alternative explanation: Employee characteristics

Another concern with my finding that training effectiveness depends on training medium and in-person training class size is that the results can be driven by differences in the type of employees who receive training in different forms, rather than difference in the level of awareness. To address this concern, I make use of the information on video training recipients. I explore two

¹⁷ I do not use the frequency of misconduct on the same topic but in a different country as a placebo because many trainings on the same topic are conducted in the same month across different countries.

employee characteristics: the relevance of their job functions to the training topic and employee level.

If the results in Tables 5 and 6 are driven by the possibility that the (small-sized) in-person training was given to employees whose job function is most relevant to the training topic, I would expect that video training given to those employees would be effective as well. To test this, I split *Video* to *Video_relevant* and *Video_irrelevant*, which is the number of employees whose job function is directly relates (or not) to the training topic. The mapping between job function and training topics is shown in Table A4 of the Internet Appendix.¹⁸ The results are shown in Table 11, Panel A. The lack of a significant coefficients on *Video_relevant* suggests that my results are not driven by the relevance of an employee's job function to the training topic.

Similarly, if the results in Tables 5 and 6 are driven by the possibility that (small-sized) in-person training was given to employees higher up in the organization, I would expect that video training given to more senior employees would be more effective. In the video training data, there is information about whether the recipient of the training is a supervisor of another employee for trainings given from January 2017. Thus, for a subsample period of April 2017 to December 2017 (because I require training information from three months prior to misconduct), I split *Video* into *Video_supervisor* and *Video_nosupervisor*, which is the number of supervisors who receive video training versus other recipients. The results are shown in Table 11 Panel B and suggest that my findings are not driven by employee level.

5.4. Addressing alternative explanation: Training content

¹⁸ I do not argue that job-topic matches that are uncolored in Table A4 indicate training that is irrelevant to the employees. For example, sales employees can engage in insider trading or employee discrimination. Rather, I assume that insider trading more directly relates to the core daily tasks of employees in the finance department than those in the sales department and thus is likely to have a greater impact.

At ManufacturingCo, training content does not differ across training medium or country. However, there are two topics (Employee relations/HR and Product) for which only one form of training was conducted. Thus, the difference in the contents according to the topic can partially drive my results. To address this concern, I re-run my analysis excluding these two topics. The results are displayed in Column (1) of Tables A5 to A8 in the Internet Appendix, and their implications are consistent with my main findings.

Further, I restrict my sample to only country-topics that have both in-person training and video training during the sample period to ensure that my results are not driven by the selection of country-topics that training was given on. Column (2) of Tables A5 to A8 in the Internet Appendix shows the results. The implications from the table are consistent with my findings in Tables 5 through 8.

5.5. Topic-by-topic analysis

In my analyses, I do not provide topic-by-topic results for two reasons. First, the number of misconduct cases for some topics is too small, disabling me to draw statistical conclusions. Second, as the contents of the training materials vary across topics, difference in the training effectiveness for each topic can be driven by the quality of the materials. However, topic-by-topic results can potentially provide some implication about the topics in which training has a larger impact. Thus, in Table A9, I show topic-by-topic results for three topics in which there were more than 50 cases of misconduct: Employee relations/ HR, Financial reporting/ Records management, and Procurement/ Bid rigging.

For the two topics on which in-person training was conducted, Financial reporting/Records management and Procurement/ Bid rigging, in-person training variables are negatively associated with the level of misconduct. However, both the economic and statistical significance is larger in

column (3), showing that Procurement/ Bid rigging is more strongly negatively related to a decrease in misconduct. This result may suggest that training is more likely to decrease misconduct which is easier to fix; for example, most bid rigging cases are committed by individuals who seek their own private benefits, in contrast to financial reporting cases which sometimes require a sophisticated planning and involve multiple individuals. However, the result could be a mere indication that the training material for Procurement/ Bid rigging is better.

5.6. Ethics training: Specificity of training content

Ethics training discusses the company's overall mission, broad values, and expected standards regarding compliance with laws. Theories predict that the specificity of compliance training contents is a desirable factor. For example, understanding what punishments or rewards follow from specific behaviors is an important aspect of learning (Bandura 1977), and high-level construals (such as ethics) have more impact on responses to distant-future events than near-future events (Trope and Liberman 2003). Therefore, I predict that ethics training will be less effective than compliance training.

To test its effect, I create the variables *In-person_Ethics* and *Video_Ethics*, which are the number of employees who receive in-person and video training on ethics in a country in a given month, and run the following regression:

$$\begin{aligned}
 Misconduct_{c,i,t} = & \alpha + \beta \cdot \sum_{t=-3}^0 In-person_{c,i,t} + \gamma \cdot \sum_{t=-3}^0 Video_{c,i,t} \\
 & + \delta \cdot \sum_{t=-3}^0 In-person_Ethics_{c,t} + \zeta \cdot \sum_{t=-3}^0 Video_Ethics_{c,t} + Fixed\ effects + \varepsilon. \quad (3)
 \end{aligned}$$

Although *In-person_Ethics* and *Video_Ethics* are country-time level variables, I still run regressions at country-topic-time level to control for the effect of compliance training on specific topics.

Table A10 in the Internet Appendix shows the results. Like previous tables, column (1) displays OLS results, and column (2) displays ZINB results. Only the coefficient on $\ln - person_Ethics_{t-3}$ in column (2) is marginally significant. The results suggest that ethics training is not effective at decreasing misconduct, as predicted.

5.7. Robustness tests

I conduct several robustness tests to show that my results are not sensitive to other methodological choices. Specifically, I conduct two tests, in addition to the zero-inflated negative binomial regressions, to address the skewed distribution of my dependent variable, *Misconduct*. As Kuvvet (2019) suggests, empirical analysis results regarding the determinants of misconduct can be driven by extreme outliers. To address this, I first run regressions using logged value of misconduct rather than the raw numbers. The results are shown in column (3) of Tables A5 to A8 in the Internet Appendix and are consistent with my main findings. Second, I run my analysis using winsorized values of training and misconduct variables. The results are shown in column (4) of Tables A5 to A8 in the Internet Appendix. The results are consistent with the main findings, except that the coefficients in Table A7 are not statistically significantly different.

Lastly, I re-run my analysis with January observations deleted. This is because, as explained in Section 3.1., local legal counsels can use their discretion to adjust the training plans. Therefore, January training may be determined by contemporaneous allegations. The results are shown in column (5) of Tables A5 to A8 in the Internet Appendix. The results are consistent with the main findings.

6. Conclusion

I investigate the effect of compliance training on corporate misconduct. Whether compliance can be taught is of empirical question. Despite having no empirical evidence, firms make large investments in it, and their legal penalty can be reduced based on it.

The difficulty in examining this research question is that data on training is proprietary, and misconduct is rarely observed publicly. I overcome this difficulty by obtaining a dataset from a large, publicly listed company. The compliance training at this firm is mandatory and determined on a three-year schedule. This helps me address omitted variables bias and reverse causality concerns.

Although I find no evidence for compliance training being effective on average, I find that training can decrease misconduct when it addresses employee attention—specifically, when training is conducted in-person. However, this effect can be mitigated in circumstances where the economic benefits (costs) of misconduct are high (low), such as in countries with higher performance pressure or weaker public enforcement. These findings provide insight into how compliance training can work and clues about how firms can better train employees, as well as giving potential guidance to regulators for reaching conclusions about the legal penalty amounts for corporate misconduct.

I caveat that my analyses may not be generalizable, as they are based on data from one firm. Nevertheless, I believe that my study provides new insights and complements other studies based on large sample size, by providing a deeper understanding on the mechanism for compliance training with detailed data. Future researchers that acquire large sample data can provide additional understanding and insights to this paper's findings.

References

- Amiram, D., Z. Bozanic, J. D. Cox, Q. Dupont, J. M. Karpoff, and R. Sloan (2018). Financial reporting fraud and other forms of misconduct: A multidisciplinary review of the literature. *Review of Accounting Studies* 23, 732–783.
- Ashbaugh-Skaife, H., D. W. Collins, and W. R. Kinney (2007). The discovery and reporting of internal control deficiencies prior to SOX-mandated audits. *Journal of Accounting and Economics* 44 (1-2), 166–192.
- Ashbaugh-Skaife, H., D. W. Collins, W. R. Kinney, and R. LaFond (2008). The effect of SOX internal control deficiencies and their remediation on accrual quality. *The Accounting Review* 83 (1), 217–250.
- Babcock, C. R. (1979) Guilty plea is entered by Lockheed. *The Washington Post*. Available at: https://www.washingtonpost.com/archive/politics/1979/06/02/guilty-plea-is-entered-by-lockheed/4a7e3a7f-6fb3-430b-aaf0-264e6e8a1703/?utm_term=.71ea03b92050
- Baloria, V. P., C. A. Marquardt, and C. I. Wiedman (2017). A lobbying approach to evaluating the whistleblower provisions of the Dodd-Frank Reform Act of 2010. *Contemporary Accounting Research* 34(3), 1305–1339.
- Banaji, M. R., M. H. Bazerman, and D. Chugh (2003). How (un)ethical are you? *Harvard Business Review* 81(12), 56–64.
- Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, NJ: Prentice Hall.
- Bandura, A., and W. Mischel (1965). Modifications of self-imposed delay of reward through exposure to live and symbolic models. *Journal of Personality and Social Psychology* 2(5), 698–705.
- Bazerman, M. H., and M. R. Banaji (2004). The social psychology of ordinary ethical failures. *Social Justice Research* 17, 111–115.
- Becker, G. S. (1968). Crime and punishment: An economic approach. In *The economic dimensions of crime*. Palgrave Macmillan, London.
- Beneish, M. D. (1999b). Incentives and penalties related to earnings overstatements that violate GAAP. *The Accounting Review* 74 (4), 425–457.
- Berger, P., and H. M. Lee (2019). Does the threat of whistleblowing reduce accounting fraud? Working paper.
- Bowen, R. M., A. C. Call, and S. Rajgopal (2010). Whistle-blowing: Target firm characteristics and economic consequences. *The Accounting Review* 85, 1239–1271.
- Call, A. C., G. S. Martin, N. Y. Sharp, and J. H. Wilde (2018). Whistleblowers and outcomes of financial misrepresentation enforcement actions. *Journal of Accounting Research* 56, 123–171.
- Chen, H., and E. Soltes (2018). Why compliance programs fail: And how to fix them. *Harvard Business Review* 96, 116–125.

- Christensen, H. B., E. Floyd, L.Y. Liu, and M. Maffett (2017). The real effects of mandated information on social responsibility in financial reports: evidence from mine-safety records. *Journal of Accounting and Economics* 64 (2-3), 284–304.
- Chugh, D., M. H. Bazerman, and M. R. Banaji (2005). Bounded ethicality as a psychological barrier to recognizing conflicts of interest. In Moore, D. A., D. M. Cain, G. F. Lowenstein, and M. H. Bazerman (Eds.), *Conflicts of interest: Challenges and solutions in business, law, medicine, and public policy* (pp. 74–95). New York, NY: Cambridge University Press.
- Cohn, S. (2012). Ex-MS Banker in China bribery case: My side of story. CNBC. Available at <https://www.cnbc.com/id/48693573>
- Committee of Sponsoring Organizations of the Treadway Commission (1992). Internal Control-Integrated Framework. Available at <https://www.coso.org/Pages/ic.aspx>
- Daske, H., L. Hail, C. Leuz, and R. Verdi (2008). Mandatory IFRS reporting around the world: Early evidence on the economic consequences. *Journal of Accounting Research* 46 (5), 1085–1142.
- Davidson, R., A. Dey, and A. Smith (2015). Executives' "off-the-job" behavior, corporate culture, and financial reporting risk. *Journal of Financial Economics* 117(1), 5–28.
- Deloitte (2016). Compliance Week. In Focus: 2016 Compliance trends survey. Available at <https://www2.deloitte.com/content/dam/Deloitte/us/Documents/governance-risk-compliance/us-advisory-compliance-week-survey.pdf>
- Dhaliwal, D., C. Hogan, R. Trezevant, and M. Wilkins (2011). Internal control disclosures, monitoring, and the cost of debt. *The Accounting Review* 86 (4), 1131–1156.
- Doyle, J., W. Ge, and S. McVay (2007b). Accruals quality and internal control over financial reporting. *The Accounting Review* 82 (5), 1141–1170.
- Dyck, A., A. Morse, and L. Zingales (2010). Who blows the whistle on corporate fraud? *The Journal of Finance* 65(6), 2213–2253.
- Dyck, A., A. Morse, and L. Zingales (2017). How pervasive is corporate fraud? Working paper.
- Ernst and Young (2016). Corporate misconduct —individual consequences. Available at [https://www.ey.com/Publication/vwLUAssets/EY-corporate-misconduct-individual-consequences/\\$FILE/EY-corporate-misconduct-individual-consequences.pdf](https://www.ey.com/Publication/vwLUAssets/EY-corporate-misconduct-individual-consequences/$FILE/EY-corporate-misconduct-individual-consequences.pdf)
- Greene, W. (2004). The behavior of the fixed effects estimator in nonlinear models. *Econometrics Journal* 7, 98–119.
- Halloran, R. (1976). Scandal costs Lockheed \$1.3 billion sale to Japan. *The New York Times*. Available at: <https://www.nytimes.com/1976/02/11/archives/scandal-costs-lockheed-13-billion-sale-to-japan-scandal-costs.html>
- Harrison, J. K. (1992). Individual and combined effects of behavior modeling and the cultural assimilator in cross-cultural management training. *Journal of Applied Psychology* 77, 952–962.

- Haugh, T. (2017). The criminalization of compliance. *Notre Dame Law Review* 92 (3), 1215–1270.
- Heese, J., and G. Pérez Cavazos (2019). When the boss comes to town: The effects of headquarters' visits on facility-level misconduct. Forthcoming at *The Accounting Review*.
- Holzman, E.R., N. T. Marshall, and B. A. Schmidt. Who's on the hot seat for an SEC investigation? Working paper.
- Iwasaki, M. (2018). Effects of external whistleblower rewards on internal reporting. Working paper.
- Karpoff, J. M., D. S. Lee, and G. S. Martin (2008a). The cost to firms of cooking the books. *Journal of Financial and Quantitative Analysis* 43, 581–611.
- Katz, D. M., and M. J. Bommarito (2014). Measuring the complexity of the law: The United States Code. *Artificial Intelligence and Law* 22(4), 337–374.
- Kedia, S., and S. Rajgopal (2011). Do the SEC's enforcement preferences affect corporate misconduct? *Journal of Accounting and Economics* 51(3), 259–278.
- Kowaleski, Z. T., A. G. Sutherland, and F. W. Vetter (2019). Can Ethics be Taught? Evidence from Securities Exams and Investment Adviser Misconduct. Working paper.
- Kuvvet, E. (2019). Are a few huge outcomes distorting financial misconduct research? *Econ Journal Watch* 16(1), 1–34.
- Lambert, D. (1992). Zero-inflated Poisson regression, with an application to defects in manufacturing. *Technometrics* 34 (1), 1–14.
- Lancaster, T. (2000). The incidental parameter problem since 1948. *Journal of Econometrics* 95, 391–413.
- Laufer, W. S. (2008). *Corporate bodies and guilty minds: The failure of corporate criminal liability*. Chicago: University of Chicago Press.
- Lazear, E.P. (2001) Educational production. *The Quarterly Journal of Economics* 116 (3), 777–803.
- Messick, D.M. (1994). Mortgage-bias complexities. *Chicago Tribune*. Available at <https://www.chicagotribune.com/news/ct-xpm-1994-03-01-9403010062-story.html>
- Miller, G. P. (2018) An economic analysis of effective compliance programs, in Jennifer Arlen (ed.) *Research handbook on corporate crime and financial misdealing*, 247-262. Cheltenham: Edward Elgar.
- Moore, D. A., P. E. Tetlock, L. Tanlu, and M. H. Bazerman (2006). Conflicts of interest and the case of auditor independence: Moral seduction and strategic issue cycling. *Academy of Management Review* 31, 10–29.

Near, J. P., and M. P. Miceli (1985). Organizational dissidence: The case of whistle-blowing. *Journal of Business Ethics* 4(1), 1–16.

Nguyen, T. (2018). The effectiveness of white-collar crime enforcement: Evidence from the war on terror. Working paper.

Podgor, S., P. J. Henning, A. Taslitz, and A. Garcia (2013). *Criminal law: Concepts and practice*, Third Edition. Durham: Carolina Academic Press.

Polinsky, A. M., and S. Shavell (1993). Should employees be subject to fines and imprisonment given the existence of corporate liability? *International Review of Law and Economics* 13(3), 239–257.

Richards, C. (1999). The transient effects of limited ethics training. *Journal of Education for Business* 74, 332–334.

Richardson, S., I. Tuna, and M. Wu (2003). Predicting earnings management: the case of earnings restatements. Working paper.

Soltes, E. (2018a). Evaluating the effectiveness of corporate compliance programs: Establishing a model for prosecutors, courts, and firms. *NYU Journal of Law and Business* 14 (3), 965–1011.

Soltes, E. (2018b). The difficulty of being good: The efficacy of integrity hotlines. Working paper.

Soltes, E. (2018c). The frequency of corporate misconduct: Public enforcement versus private reality. Working paper.

Stanford Securities Class Action Clearinghouse (SSCAC) (2019). Security class action filings. Available at: <http://securities.stanford.edu/research-reports/1996-2018/Cornerstone-Research-Securities-Class-Action-Filings-2018-YIR.pdf>

Stansbury, J., and B. Barry. (2007). Ethics programs and the paradox of control. *Business Ethics Quarterly*, 17: 239–261.

Stubben, S. R., and K. T. Welch (2018). Evidence on the use and efficacy of internal whistleblowing systems. Working paper.

Stubben, S. R., and K. T. Welch (2019). Are Secondhand Internal Whistleblowing Reports Credible? Working paper.

Technavio (2018). Global Corporate Compliance Training Market 2017-2021. Available at https://www.technavio.com/report/global-k12-and-higher-education-global-corporate-compliance-training-market-2017-2021&utm_source=T3&utm_medium=BW&utm_campaign=Media

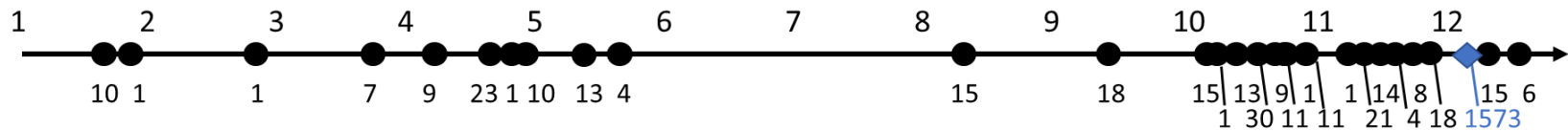
Tenbrunsel, A. E., and D. M. Messick (2004). Ethical fading: The role of self deception in unethical behavior. *Social Justice Research* 17(2), 223–236.

- Thomson Reuters (2018). Cost of compliance 2018. Available at <https://risk.thomsonreuters.com/content/dam/openweb/documents/pdf/risk/report/cost-of-compliance-special-report-2018.pdf>
- Trope, Y., and N. Liberman (2003). Temporal construal. *Psychological Review* 95, 3–15.
- USSC (2004). Federal Sentencing Guidelines for Organizations. Available at <https://www.ussc.gov/guidelines/2018-guidelines-manual/2018-chapter-8>
- Warren, D. E., J. P. Gaspar, and W. S. Laufer (2014). Is formal ethics training merely cosmetic? A study of ethics training and ethical organizational culture. *Business Ethics Quarterly* 24, 85–117.
- Weaver, G. R., and L. K. Treviño (1999). Compliance and values oriented ethics programs: Influences on employees' attitudes and behavior. *Business Ethics Quarterly* 9, 315–335.
- Wiedman, C., and C. Zhu (2018). Do the SEC whistleblower provisions of Dodd Frank deter aggressive financial reporting? Working paper.
- Wooldridge, J.M. (2002). *Econometric analysis of cross section and panel data* (first edition). The MIT Press, Cambridge.

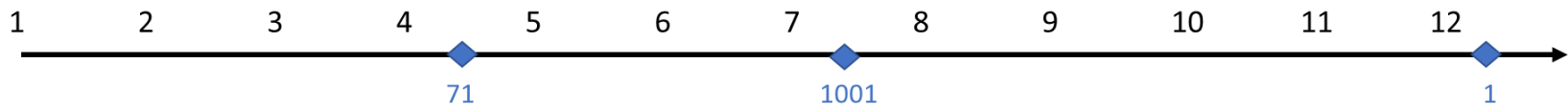
Figure 1. Examples of training

This figure shows three examples of how compliance training was conducted in a country-year on a topic. Each black dot represents an in-person training, and each blue diamond represents a video training. The numbers below the dots (diamonds) represent the number of employees who received each training. The numbers above the arrow indicate the start of months, to show when the trainings were conducted. For example, a dot between 1 and 2 (noted above the arrow) represents a training that occurred in January.

Example 1. Country= United Kingdom, Topic= Antitrust/Sales, Year= 2015



Example 2. Country= United States, Topic= Bribery/Government relations, Year= 2014



Example 3. Country= Mexico, Topic= Security, Year= 2016

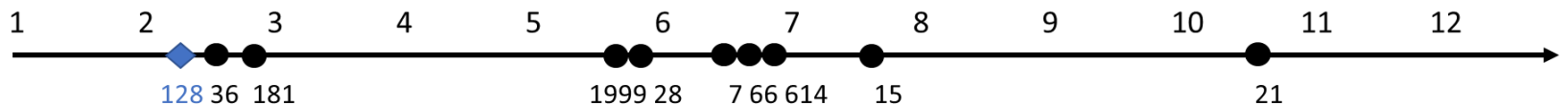
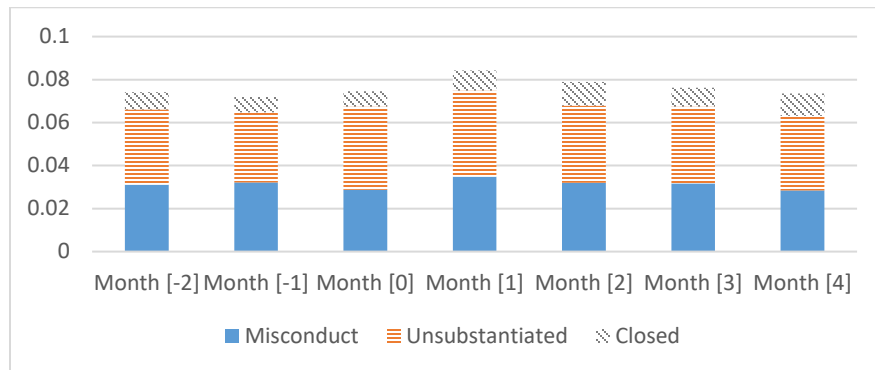


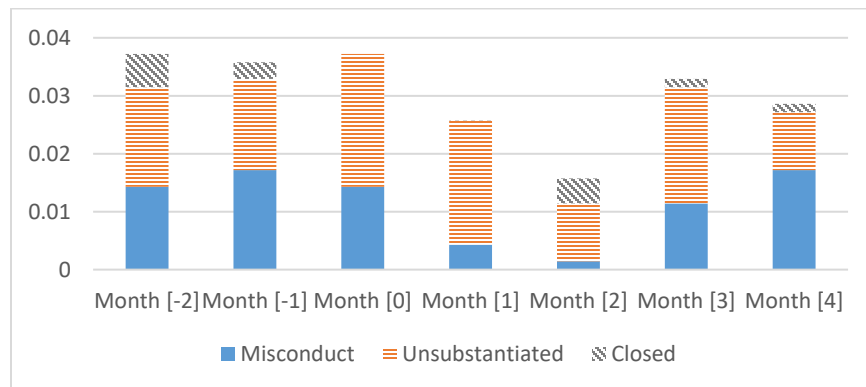
Figure 2. Number of allegations around training

This figure shows the average frequency of misconduct, unsubstantiated allegations, and allegations closed without conclusion around the month that compliance training was given. Panel A includes all training events, regardless of medium; Panel B includes in-person training; and Panel C includes video training.

Panel A. All training



Panel B. In-person training



Panel C. Video training

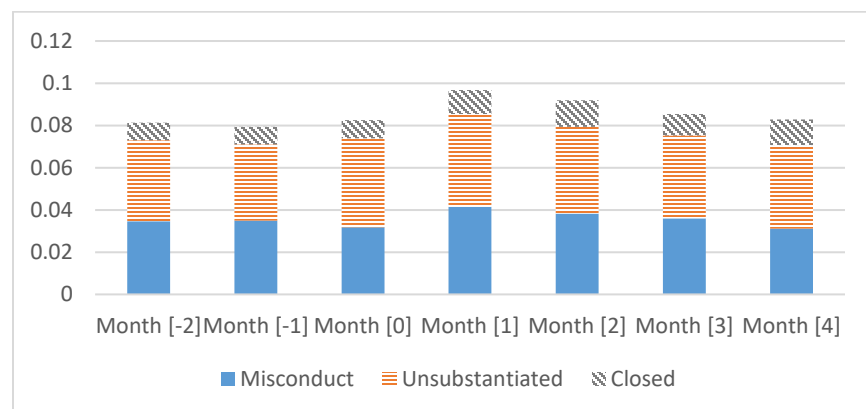


Figure 3. The two effects of compliance training and reporting probability

This figure shows the expected relation between the two effects of compliance training (the knowledge effect and the awareness effect) and reporting probability of misconduct (substantiated allegations) and unsubstantiated allegations.

	Misconduct reporting	Unsubstantiated allegations reporting
Knowledge effect	↑	↓
Awareness effect	↑	↑
Aggregate effect	↑	?

Table 1. Frequency of allegations and training participants by topic

This table reports the number of allegations (columns (1) through (4)), the number of employees who received in-person compliance training (columns (5) and (6)), and the number of employees who received video training (columns (7) and (8)), by topic. The sample period is from January 2013 to December 2017.

Topic	Allegations				In-person participants		Video participants	
	(1) Total	(2) (%)	(3) Misconduct	(4) Unsubstan- -tiated	(5) Total	(6) (%)	(7) Total	(8) (%)
Compliance training								
Antitrust	16	0.7%	7	9	16,513	35.1%	62,291	13.8%
Bribery/Government relations	47	2.1%	23	21	5,771	12.3%	77,251	17.1%
Employee relations/HR	1,698	75.9%	709	845	0	0.0%	38,484	8.5%
Environment	5	0.2%	2	3	0	0.0%	0	0.0%
Financial reporting/Records management	180	8.0%	105	60	1,517	3.2%	52,500	11.6%
Insider trading	8	0.4%	5	3	1,222	2.6%	19,456	4.3%
Procurement/ Bid rigging	179	8.0%	72	80	2,595	5.5%	31,050	6.9%
Product	49	2.2%	16	21	510	1.1%	0	0.0%
Security	55	2.5%	18	33	4,903	10.4%	10,971	2.4%
Total	2,237	100.0%	957	1,075	33,031	70.3%	292,003	64.5%
Ethics training					13,961	29.7%	160,423	35.5%
Total	2,237	100%	957	1,075	46,992	100.0%	452,426	100.0%

Table 2. Frequency of allegations and training participants by region and year

This table reports the number of allegations (columns (1) through (3)), the number of employees who received in-person compliance training (columns (4) and (5)), and the number of employees who received video training (columns (6) and (7)), by region (Panel A) and by year (Panel B). The sample period is from January 2013 to December 2017.

Panel A. Frequency of allegations and training participants by region

Region	No. of country	% of employee population	Allegations			In-person participants		Video participants	
			(1) Total	(2) Misconduct	(3) Unsubstantiated	(4) Compliance	(5) Ethics	(6) Compliance	(7) Ethics
Asia, Middle East and Africa	30	29%	414	171	175	8,008	8,410	81,510	52,618
Europe	37	32%	174	82	73	8,563	1,867	102,585	55,981
Latin America	19	24%	967	516	415	14,898	3,115	49,427	26,529
North America	2	15%	682	188	412	1,562	569	58,481	25,295
Total	88	100%	2,237	957	1,075	33,031	13,961	292,003	160,423

Panel B. Frequency of allegations and training participants by year

Year	Allegations			In-person participants		Video participants	
	(1) Total	(2) Misconduct	(3) Unsubstantiated	(4) Compliance	(5) Ethics	(6) Compliance	(7) Ethics
2013	369	144	193	4,156	2,531	64,300	45,106
2014	465	197	220	7,927	2,360	32,077	58,679
2015	430	175	220	3,326	4,049	57,662	14,160
2016	507	244	223	8,345	1,332	45,252	42,029
2017	466	197	219	9,277	3,689	92,712	449
Total	2,237	957	1,075	33,031	13,961	292,003	160,423

Table 3. Descriptive statistics

This table shows the summary statistics for the variables used in my analyses. *Allegation* is the number of allegations on a topic in a country in a month. *Misconduct (Unsubstantiated)* is the number of substantiated (unsubstantiated) allegations on a topic in a country in a month. *Train (In-person, Video)* is the number of employees who received (in-person, video) compliance training on a topic in a country in a month. *In-person_small (In-person_big)* is the number of employees who received in-person compliance training on a topic in a class less than or equal to (more than) 12 people.

	N	Mean	Median	StD	Min	Max
<i>Allegation</i>	38,760	0.06	0	0.57	0	24
<i>Misconduct</i>	38,760	0.02	0	0.26	0	10
<i>Unsubstantiated</i>	38,760	0.03	0	0.34	0	15
<i>Train</i>	38,760	8.37	0	99.90	0	6,449
<i>In-person</i>	38,760	0.84	0	22.00	0	2,080
<i>In-person_small</i>	38,760	0.08	0	0.93	0	33
<i>In-person_big</i>	38,760	0.76	0	21.95	0	2,080
<i>Video</i>	38,760	7.53	0	96.88	0	6,449

Table 4. Compliance training and misconduct

This table presents OLS and zero-inflated negative binomial regression results that tests the effect of compliance training on misconduct. Variables are defined in the header of Table 3. The dependent variable is the number of substantiated allegations. T-statistics for OLS are based on standard errors clustered by country-topic and by time and those for ZINB are based on robust standard errors. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively (two-tailed).

	(1)	(2)	(3)	(4)
	<i>Misconduct</i>	<i>Misconduct</i>	<i>Misconduct</i>	<i>Misconduct</i>
	(OLS)	(OLS)	(OLS)	(ZINB)
$Train_t$ (thou)	-0.016 (-0.745)	-0.028 (-1.359)	-0.065 (-1.411)	-0.083 (-0.543)
$Train_{t-1}$ (thou)	0.071 (0.706)	0.059 (0.636)	0.023 (0.411)	0.212 (1.365)
$Train_{t-2}$ (thou)	0.038 (0.649)	0.028 (0.473)	-0.008 (-0.187)	0.162 (0.923)
$Train_{t-3}$ (thou)	0.026 (0.545)	0.016 (0.535)	-0.022 (-1.622)	0.042 (0.371)
<i>Employees</i> (thou)	0.043** (2.634)			
Observations	38,760	38,760	38,760	38,760
(Pseudo) R-squared	0.077	0.107	0.458	0.394
Country FE	No	Yes	No	
Topic FE	Yes	Yes	No	
Country-Topic FE	No	No	Yes	
Time FE	Yes	Yes	Yes	
Inflator				Country-Topic, Time

Table 5. Compliance training and misconduct by training medium

This table presents OLS and zero-inflated negative binomial regression results that test the effect of compliance training on misconduct by training medium. Variables are defined in the header of Table 3. The dependent variable in Panel A is the number of substantiated allegations. T-statistics for OLS are based on standard errors clustered by country-topic and by time and those for ZINB are based on robust standard errors. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively (two-tailed).

	(1) <i>Misconduct</i> (OLS)	(2) <i>Misconduct</i> (ZINB)
$\ln - person_t$ (thou)	0.015 (0.902)	-1.701 (-0.971)
$\ln - person_{t-1}$ (thou)	-0.028** (-2.140)	-158.441** (-2.092)
$\ln - person_{t-2}$ (thou)	-0.024* (-1.674)	-1,667.105* (-1.958)
$\ln - person_{t-3}$ (thou)	-0.007 (-0.726)	-8.184* (-1.750)
$Video_t$ (thou)	-0.070 (-1.450)	-0.081 (-0.540)
$Video_{t-1}$ (thou)	0.026 (0.431)	0.207 (1.345)
$Video_{t-2}$ (thou)	-0.008 (-0.160)	0.158 (0.912)
$Video_{t-3}$ (thou)	-0.022 (-1.589)	0.035 (0.312)
Observations	38,760	38,760
(Pseudo) R-squared	0.458	0.400
Country-Topic FE	Yes	
Time FE	Yes	
Inflator		Country-Topic, Time

Table 6. In-person compliance training class size and misconduct

This table presents OLS and zero-inflated negative binomial regression results that test the effect of in-person compliance training in different class size. Variables are defined in the header of Table 3. T-statistics for OLS are based on standard errors clustered by country-topic and by time and those for ZINB are based on robust standard errors. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively (two-tailed).

	(1) <i>Misconduct</i> (OLS)	(2) <i>Misconduct</i> (ZINB)
$\ln - person_small_t$ (thou)	-0.642 (-1.362)	-143.333* (-1.698)
$\ln - person_small_{t-1}$ (thou)	-0.792** (-2.029)	-489.810** (-2.026)
$\ln - person_small_{t-2}$ (thou)	-0.810* (-1.684)	-1,634.852* (-1.913)
$\ln - person_small_{t-3}$ (thou)	-0.910** (-2.028)	-193.171 (-1.474)
$\ln - person_big_t$ (thou)	0.017 (0.957)	-1.166 (-0.811)
$\ln - person_big_{t-1}$ (thou)	-0.025** (-2.106)	-108.386** (-2.303)
$\ln - person_big_{t-2}$ (thou)	-0.021 (-1.623)	-384.469*** (-12.224)
$\ln - person_big_{t-3}$ (thou)	-0.004 (-0.436)	-6.561* (-1.779)
Observations	38,760	38,760
(Pseudo) R-squared	0.458	0.401
Video training controls	Yes	Yes
Country-Topic FE	Yes	
Time FE	Yes	
Inflator		Country-Topic, Time

Table 7. Compliance training and misconduct under different levels of performance pressure

This table presents OLS regression results that test the effect of compliance training under different levels of performance pressure. Performance pressure is measured as whether production facilities in a country met their efficiency targets during the sample period. Production efficiency is calculated as net operating time divided by total hours used. Variables are defined in the header of Table 3. T-statistics are based on standard errors clustered by country-topic and by time. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively (two-tailed).

	(1)	(2)	
	Production target		
	Miss	Meet	
	<i>Misconduct</i>	<i>Misconduct</i>	$\beta_{miss} \neq \beta_{meet}$
	(OLS)	(OLS)	p-value
<i>In – person_t (thou)</i>	0.022	0.150	0.631
	(0.967)	(0.582)	
<i>In – person_{t-1} (thou)</i>	-0.029*	-0.181*	0.154
	(-1.878)	(-1.826)	
<i>In – person_{t-2} (thou)</i>	0.000	-0.128*	0.061*
	(0.043)	(-1.960)	
<i>In – person_{t-3} (thou)</i>	-0.021	-0.061	0.479
	(-0.716)	(-1.272)	
<i>Video_t (thou)</i>	-0.053	-0.022	0.472
	(-1.049)	(-0.616)	
<i>Video_{t-1} (thou)</i>	0.087	0.108	0.754
	(0.907)	(1.159)	
<i>Video_{t-2} (thou)</i>	-0.032	-0.032*	0.994
	(-0.409)	(-1.732)	
<i>Video_{t-3} (thou)</i>	-0.070***	0.038	0.006***
	(-2.966)	(0.991)	
Observations	8,032	7,032	
(Pseudo) R-squared	0.579	0.114	
Country-Topic FE	Yes	Yes	
Time FE	Yes	Yes	

Table 8. Compliance training and misconduct under different public enforcement environment

This table presents OLS regression results that test the effect of compliance training in different public enforcement environments. Public enforcement is defined as the Rule of Law index published by the World Bank. Strong (weak) public enforcement includes countries whose Rule of Law index is above (below or equal to) the median. Variables are defined in the header of Table 3. T-statistics are based on standard errors clustered by country-topic and by time. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively (two-tailed).

	(1)	(2)	
	Enforcement		
	Strong	Weak	
	<i>Misconduct</i>	<i>Misconduct</i>	$\beta_{strong} \neq \beta_{weak}$
	(OLS)	(OLS)	p-value
$ln - person_t (thou)$	0.213 (0.747)	-0.011 (-1.016)	0.450
$ln - person_{t-1} (thou)$	-0.138** (-2.274)	-0.028 (-1.664)	0.054*
$ln - person_{t-2} (thou)$	-0.128** (-2.573)	-0.024 (-1.666)	0.057*
$ln - person_{t-3} (thou)$	-0.065 (-1.085)	0.001 (0.111)	0.236
$Video_t (thou)$	-0.091 (-1.367)	-0.008 (-0.208)	0.343
$Video_{t-1} (thou)$	-0.008 (-0.290)	0.125 (1.092)	0.014**
$Video_{t-2} (thou)$	-0.045 (-1.078)	0.100 (0.992)	0.123
$Video_{t-3} (thou)$	-0.022 (-0.903)	-0.021 (-0.403)	0.944
Observations	19,152	19,608	
(Pseudo) R-squared	0.448	0.466	
Country-Topic FE	Yes	Yes	
Time FE	Yes	Yes	

Table 9. Probability of reporting

This table addresses concerns about the dependent variable (*Misconduct*) regarding the probability of misconduct being reported. In column (1), the dependent variable is *Unsubstantiated*, which is the number of unsubstantiated allegations. In column (2), the dependent variable is *Misconduct (excluding Reporting Up)*, which is the number of substantiated allegations, excluding those made through the internal “reporting up” system. In column (3), two topics that are most likely to be alleged by the employees (employee relations/HR and security) are dropped from the sample. Variables are defined in the header of Table 3. T-statistics are based on standard errors clustered by country-topic and by time. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively (two-tailed).

	(1)	(2)	(3)
	<i>Unsubstantiated</i>	<i>Misconduct</i> <i>(excluding Reporting Up)</i>	<i>Misconduct</i>
	(OLS)	(OLS)	(OLS)
<i>In – person_t (thou)</i>	0.103 (1.068)	0.017 (0.902)	-0.009 (-0.992)
<i>In – person_{t-1} (thou)</i>	0.010 (0.384)	-0.021** (-2.117)	-0.029* (-1.729)
<i>In – person_{t-2} (thou)</i>	-0.002 (-0.076)	-0.018* (-1.733)	-0.028 (-1.474)
<i>In – person_{t-3} (thou)</i>	0.019 (1.399)	-0.007 (-0.614)	-0.008 (-0.697)
<i>Video_t (thou)</i>	0.028 (1.000)	-0.050 (-1.514)	0.006 (0.546)
<i>Video_{t-1} (thou)</i>	-0.003 (-0.091)	0.017 (0.419)	-0.004 (-0.822)
<i>Video_{t-2} (thou)</i>	0.068 (1.322)	-0.004 (-0.083)	0.003 (0.323)
<i>Video_{t-3} (thou)</i>	-0.127 (-1.409)	-0.028 (-1.001)	0.015 (1.048)
Observations	38,760	38,760	29,070
R-squared	0.622	0.461	0.096
Country-Topic FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes

Table 10. Placebo tests and reverse causality test

This table presents OLS results of the placebo and reverse causality tests. Variables are defined in the header of Table 3. The dependent variable in column (1) is the number of substantiated allegations on topics other than the topic of interest in a given country in a month. That in column (2) is the number of substantiated allegations four months after the training, and that in (3) is the number of substantiated allegations in the month of training. T-statistics for OLS are based on standard errors clustered by country-topic and by time. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively (two-tailed).

	(1) <i>Other Misconduct</i> (OLS)	(2) <i>Misconduct_{t+4}</i> (OLS)	(3) <i>Misconduct</i> (OLS)
<i>In – person_{t+2} (thou)</i>			0.013 (0.412)
<i>In – person_{t+1} (thou)</i>			0.052 (0.752)
<i>In – person_t (thou)</i>	0.521 (1.145)	0.000 (0.003)	0.013 (0.597)
<i>In – person_{t-1} (thou)</i>	-0.051 (-0.227)	0.073 (1.049)	-0.022** (-2.067)
<i>In – person_{t-2} (thou)</i>	0.487 (1.626)	-0.015 (-0.840)	-0.037 (-1.383)
<i>In – person_{t-3} (thou)</i>	-0.232 (-1.009)	-0.020 (-1.147)	-0.015 (-0.895)
<i>Video_{t+2} (thou)</i>			-0.044 (-1.188)
<i>Video_{t+1} (thou)</i>			0.050 (1.270)
<i>Video_t (thou)</i>	-0.015 (-0.103)	0.022 (1.165)	-0.074 (-1.416)
<i>Video_{t-1} (thou)</i>	0.164 (1.310)	-0.042 (-0.792)	0.024 (0.399)
<i>Video_{t-2} (thou)</i>	-0.066 (-1.644)	-0.119 (-1.503)	-0.009 (-0.171)
<i>Video_{t-3} (thou)</i>	-0.088 (-0.909)	0.098 (1.396)	-0.028 (-1.281)
Observations	38,760	35,360	37,400
R-squared	0.486	0.490	0.467
Country-Topic FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes

Table 11. Characteristics of video training recipients and misconduct

This table presents OLS and zero-inflated negative binomial regression results that test the differential effect of video compliance training, according to employee characteristics. Panel A examines the relevance of employee job function to the topic of misconduct, and Panel B examines employee level. *Video_relevant* (*Video_irrelevant*) is the number of employees who received video training among those whose job function is more related (less related) to the training topic. Relevance is defined in Table A4 of the internet appendix. *Video_supervisor* (*Video_nosupervisor*) is the number of supervisors (employees other than supervisors) who received video training. T-statistics are based on standard errors clustered by country-topic and by time. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively (two-tailed).

Panel A. Relevance of employee job function to the topic of misconduct

	(1) <i>Misconduct</i> (OLS)	(2) <i>Misconduct</i> (ZINB)
<i>Video_relevant_t</i> (thou)	0.034 (1.304)	-7.642 (-0.633)
<i>Video_relevant_{t-1}</i> (thou)	-0.010 (-0.256)	-0.842 (-0.850)
<i>Video_relevant_{t-2}</i> (thou)	-0.003 (-0.107)	-1.270 (-0.743)
<i>Video_relevant_{t-3}</i> (thou)	0.005 (0.320)	-5.972 (-0.329)
<i>Video_irrelevant_t</i> (thou)	-0.094 (-1.646)	0.284 (0.423)
<i>Video_irrelevant_{t-1}</i> (thou)	0.036 (0.437)	0.578 (0.439)
<i>Video_irrelevant_{t-2}</i> (thou)	-0.009 (-0.133)	0.554 (0.483)
<i>Video_irrelevant_{t-3}</i> (thou)	-0.029 (-1.653)	0.285 (1.139)
Observations	38,760	38,760
(Pseudo) R-squared	0.459	0.401
In-person training controls	Yes	Yes
Country-Topic FE	Yes	
Time FE	Yes	
Inflator		Country-Topic, Time

Panel B. Employee level

	(1) <i>Misconduct</i> (OLS)	(2) <i>Misconduct</i> (ZINB)
<i>Video_supervisor_t (thou)</i>	-0.066 (-0.315)	1.787 (0.501)
<i>Video_supervisor_{t-1} (thou)</i>	-0.823 (-1.337)	-1.508 (-0.450)
<i>Video_supervisor_{t-2} (thou)</i>	-0.190 (-1.459)	2.584 (0.196)
<i>Video_supervisor_{t-3} (thou)</i>	-0.557 (-1.794)	-1.967 (-0.727)
<i>Video_supervisor_t (thou)</i>	-0.149 (-0.924)	-0.306 (-0.418)
<i>Video_supervisor_{t-1} (thou)</i>	0.453 (1.252)	0.985 (1.460)
<i>Video_supervisor_{t-2} (thou)</i>	0.052 (0.915)	-0.203 (-0.061)
<i>Video_supervisor_{t-3} (thou)</i>	0.193 (1.377)	0.864 (1.462)
Observations	6,120	6,120
(Pseudo) R-squared	0.557	0.569
In-person training controls	Yes	Yes
Country-Topic FE	Yes	
Time FE	Yes	
Inflator		Country-Topic, Time

Appendix 1. Examples of allegations by topic

Antitrust

“An anonymous allegation that the Director of Sales for *Country A* is participating in fixing market prices ...”

Bribery/Government relations

“General allegation that Plant Manager and local management are providing benefits to local politicians directly or through vendors, including awarding a contract to a vendor that is fronting for the nephew of a local politician. We have emailed the reporter to meet with us and provide us details.”

Employee relations/HR

“Allegation against HR Director for making discriminatory (sexual orientation and gender) and inappropriate remarks to employees.”

“An allegation of discriminatory hiring practices at *Plant B*. The reporter alleges 96% of the employees are *Race C*. The reporter also threatens to call the Equal Employment Opportunity Commission and Civil Liberty’s Union and plans to start a boycott with *Product D*.”

Environment

“An anonymous report that management is directing the Caller to underreport the data for landfill waste.”

Financial reporting/Records management

“An anonymous report alleging the Regional Supervisor and the Country Manager are falsely overselling in order to reach sales goals. The Caller knows this is happening ... with *Products E and F*.”

“Salesperson alleges that his supervisors are manipulating reported sales numbers and that they are threatening him for exposing this issue.”

Insider trading

“Employee had set up an automatic order to exercise some share options if the price went above certain amount ... within the earnings-related blackout period.”

Procurement/ Bid rigging

“Allegation that Engineering Officer is involved in receiving payments from *vendors G* in exchange for issuing them purchase orders.”

Product

“A consumer found [*a sharp object*] in *Product H* ...”

Security

“An anonymous allegation that the company’s ‘golden rule’ of not putting your hands on the machines while they’re running is not being followed.”

Appendix 2. Average penalty amount per topic

This table reports the average penalty amounts per topic as reported in Violation Tracker database. This database includes enforcement data obtained from more than 40 different federal regulatory agencies and all divisions of the Justice Department and selected types of class-action lawsuits since 2000. Each violation is classified into a corresponding topic, based on the primary offense information in the database. Calculation is based on data violations from January 2000 to December 2018.

Topic	Primary Offense (Violation Tracker)	Average penalty amount (Violation Tracker)	Median penalty amount (Violation Tracker)
Antitrust	Price-fixing or anti-competitive practices	\$45,584,293	\$12,000,000
Bribery/Government relations	Foreign Corrupt Practices Act Kickbacks and bribery	\$42,868,731	\$10,145,000
Employee relations/HR	Americans with Disabilities Act	\$274,818	\$14,793
	Benefit plan administrator violation		
	Child labor or youth employment violation		
	Discriminatory practices		
	Employment discrimination		
	Employment screening violation		
	Fair Labor Standards Act		
	Family and Medical Leave Act		
	Labor relations violation		
	Payday lending violation		
Environment	Wage and hour violation	\$2,301,926	\$17,600
	Work visa violations		
Financial reporting/Records management	Environmental violation	\$21,346,384	\$2,050,000
	Accounting fraud or deficiencies		
Insider trading	Data submission deficiencies	\$2,000,000	\$2,000,000
Procurement/ Bid rigging	Insider trading*		
Product	Bid rigging**	\$4,548,359	\$432,598
Security	Product safety violation	\$942,395	\$345,000
	Workplace safety or health violation	\$15,759	\$8,209

* Insider trading penalty amount is calculated excluding violations incurred by financial institutions, because the nature and the scale of insider trading misconduct can differ between financial institutions and others.

** Bid rigging cases are recorded as a secondary offense in Violation Tracker.

Internet Appendix for

Does Compliance Training Decrease Corporate Misconduct? Evidence from Field Data

Jihwon Park

Harvard Business School

January 2020

Table A1. Compliance training and misconduct under different levels of performance pressure: Zero-inflated negative binomial (ZINB)

This table presents ZINB regression results that tests the effect of compliance training under different levels of performance pressure. Performance pressure is measured as whether production facilities in a country met its efficiency target during the sample period. Production efficiency is calculated as net operating time divided by total hours used. T-statistics are based on robust standard errors. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively (two-tailed).

	(1)	(2)	
	Production target		
	Miss	Meet	
	<i>Substantiated</i>	<i>Substantiated</i>	$\beta_{miss} \neq \beta_{meet}$
	(ZINB)	(ZINB)	p-value
<i>In – person_t (thou)</i>	-2.820** (-2.035)	43.304 (0.180)	0.002***
<i>In – person_{t-1} (thou)</i>	-341.938*** (-3.824)	-59.307 (-0.289)	0.755
<i>In – person_{t-2} (thou)</i>	-853.711*** (-2.648)	-7,712.153*** (-5.109)	0.024**
<i>In – person_{t-3} (thou)</i>	-9.356*** (-2.919)	-58.915 (-0.085)	0.629
<i>Video_t (thou)</i>	-0.020 (-0.130)	1.450 (0.073)	0.904
<i>Video_{t-1} (thou)</i>	0.288*** (3.963)	1.240 (0.238)	0.045**
<i>Video_{t-2} (thou)</i>	0.064 (0.421)	-9.186 (-0.179)	0.112
<i>Video_{t-3} (thou)</i>	0.010 (0.106)	0.447 (0.410)	0.534
Observations	8,032	7,032	
Pseudo R-squared	0.438	0.465	
Inflator	Country-Topic, Time	Country-Topic, Time	

Table A2. Compliance training and misconduct under different levels of performance pressure: Using employee turnover as proxy for performance pressure

This table presents OLS regression results that tests the effect of compliance training under different levels of performance pressure. Performance pressure is measured by the employee turnover ratio, calculated as the number of employees that left the firm divided by the total number of employees in a given country-year. Turnover is classified as ‘High’ (‘Low’) if the turnover rate is higher than (lower than or equal to) the median. T-statistics are based on standard errors clustered by country-topic and by time. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively (two-tailed).

	(1)	(2)	
	Turnover		
	High	Low	
	<i>Substantiated</i>	<i>Substantiated</i>	$\beta_{high} \neq \beta_{low}$
	(OLS)	(OLS)	p-value
$In - person_t (thou)$	0.026 (1.239)	-0.027** (-2.298)	0.007***
$In - person_{t-1} (thou)$	-0.029** (-2.189)	-0.023** (-2.595)	0.523
$In - person_{t-2} (thou)$	-0.016 (-1.143)	-0.030** (-2.230)	0.685
$In - person_{t-3} (thou)$	0.005 (0.369)	-0.020** (-2.337)	0.037**
$Video_t (thou)$	-0.057 (-1.154)	-0.012 (-0.608)	0.471
$Video_{t-1} (thou)$	0.104 (1.040)	-0.035** (-2.257)	0.164
$Video_{t-2} (thou)$	-0.034 (-0.526)	0.032 (1.406)	0.388
$Video_{t-3} (thou)$	-0.054** (-2.414)	0.033 (0.978)	0.08*
Observations	17,184	17,304	
R-squared	0.538	0.231	
Country-Topic FE	Yes	Yes	
Time FE	Yes	Yes	

Table A3. Compliance training and misconduct under different public enforcement environment: Zero-inflated negative binomial (ZINB)

This table presents ZINB regression results that tests the effect of compliance training in different public enforcement environments. Public enforcement is defined as Rule of Law index published by the World Bank. Strong (weak) public enforcement includes countries whose Rule of Law index is above (below or equal to) the median. T-statistics are based on robust standard errors. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively (two-tailed).

	(1)	(2)	
	Enforcement		
	Strong	Weak	
	<i>Substantiated</i>	<i>Substantiated</i>	$\beta_{strong} \neq \beta_{weak}$
	(ZINB)	(ZINB)	p-value
$\ln - person_t (thou)$	-0.860 (-0.899)	-4.490 (-0.686)	0.915
$\ln - person_{t-1} (thou)$	-363.073*** (-3.910)	-125.491** (-1.994)	0.023**
$\ln - person_{t-2} (thou)$	-476.277** (-1.983)	-10,121.951*** (-10.769)	0.693
$\ln - person_{t-3} (thou)$	-7.772 (-1.487)	-7.178 (-1.524)	0.997
$Video_t (thou)$	-0.046 (-0.268)	-0.074 (-0.475)	0.994
$Video_{t-1} (thou)$	0.191 (1.126)	0.435 (1.394)	0.983
$Video_{t-2} (thou)$	0.066 (0.449)	0.859** (2.182)	0.885
$Video_{t-3} (thou)$	0.104 (1.064)	-0.479 (-1.644)	0.968
Observations	19,152	19,608	
Pseudo R-squared	0.438	0.402	
Inflator	Country-Topic, Time	Country-Topic, Time	

Table A4. Job function relevance

This table tabulates the relation between video training topics and employees' job functions. Each row represents video training topics, and each column represents employee's job functions. The numbers in each cell is the number of participants of corresponding training topics from each job function. The colored cells are defined to be the cases in which video training is given to employees whose job function is relevant to the topic.

	Sales	Corporate/ Government Affairs	Finance	Procurement	Supply Chain Management	HR	Enginee ring	Manufac turing	Customer Service, Logistics
Antitrust/ Sales	31,244	349	3,627	2,086	636	2,036	720	5,499	5,264
Bribery/Government relations	27,384	410	5,502	2,358	1,029	3,257	1,270	13,567	8,309
Financial reporting/Records management	19,615	299	3,501	1,457	537	2,346	816	8,360	5,328
Insider trading	4,778	227	1,873	808	427	1,147	531	1,969	1,739
Procurement/ Bid rigging	10,913	152	2,349	1,189	375	1,338	527	5,497	3,365
Employee relations/HR	16,596	194	1,971	896	450	1,309	554	5,580	4,170
Security	2,520	136	1,027	437	221	692	286	1,211	960
Total	113,050	1,767	19,850	9,231	3,675	12,125	4,704	41,683	29,135

	General Management	Information Technology, Solutions	Legal, Business Integrity, Security	Marketing	RDQ	Shared Services	Strategy, Insights Analytics	No Function	Total
Antitrust/ Sales	224	1,368	605	3,304	3,600	994	524	211	62,291
Bribery/Government relations	222	1,801	603	3,237	5,392	2,056	513	341	77,251
Financial reporting/Records management	141	1,257	367	2,345	3,420	2,000	369	342	52,500
Insider trading	139	1,187	292	1,458	1,999	367	373	142	19,456
Procurement/ Bid rigging	99	677	203	1,313	1,997	745	222	89	31,050
Employee relations/HR	83	993	286	1,332	2,831	977	234	28	38,484
Security	109	632	253	891	1,035	191	213	157	10,971
Total	1,017	7,915	2,609	13,880	20,274	7,330	2,448	1,310	292,003

Table A5. Robustness test: Compliance training and misconduct by training medium

This table presents robustness test results for Table 5. Column (1) shows results excluding Employee relations/HR and Product observations. Column (2) shows result with only the country-topic matches for which both in-person training and video training was conducted during the sample period. Column (3) shows results with $\log(1+Misconduct)$ as the dependent variable. Column (4) shows results with training and misconduct variables winsorized at 1% level (top and bottom 0.5%). Column (5) shows results excluding January observations. T-statistics are based on standard errors clustered by country*topic and by time. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively (two-tailed).

	(1) <i>Misconduct</i> (OLS)	(2) <i>Misconduct</i> (OLS)	(3) $\log(Misconduct)$ (OLS)	(4) <i>Misconduct</i> (OLS)	(5) <i>Misconduct</i> (OLS)
$In - person_t (thou)$	0.012 (0.684)	0.013 (0.735)	0.009 (0.812)	0.010 (0.505)	0.012 (0.716)
$In - person_{t-1} (thou)$	-0.023* (-1.945)	-0.024** (-2.119)	-0.018** (-2.137)	-0.031** (-2.037)	-0.028** (-2.028)
$In - person_{t-2} (thou)$	-0.021 (-1.506)	-0.021 (-1.499)	-0.016* (-1.682)	-0.030* (-1.771)	-0.029** (-2.065)
$In - person_{t-3} (thou)$	-0.006 (-0.731)	-0.005 (-0.620)	-0.005 (-0.838)	-0.008 (-0.738)	-0.006 (-0.604)
$Video_t (thou)$	0.004 (0.421)	0.001 (0.047)	-0.025 (-1.589)	-0.005 (-0.500)	-0.065 (-1.519)
$Video_{t-1} (thou)$	-0.005 (-1.007)	-0.013 (-1.669)	0.003 (0.131)	0.014 (0.663)	0.033 (0.508)
$Video_{t-2} (thou)$	0.001 (0.159)	0.004 (0.342)	-0.007 (-0.445)	0.008 (0.457)	0.055 (1.423)
$Video_{t-3} (thou)$	0.013 (0.918)	-0.006 (-0.429)	-0.000 (-0.093)	0.007 (0.636)	-0.020 (-1.602)
Observations	29,070	0.013	38,760	38,760	36,040
R-squared	0.095	(0.735)	0.445	0.307	0.463
Robustness test type	Exclude HR and Product	Include country- topics w/ both training mediums	Log	Winsor	Exclude January
Country-Topic FE	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes

Table A6. Robustness test: In-person compliance training group size and misconduct

This table presents robustness test results for Table 6. Column (1) shows results excluding Employee relations/HR and Product observations. Column (2) shows result with only the country-topic matches for which both in-person training and video training was conducted during the sample period. Column (3) shows results with $\log(1+Misconduct)$ as the dependent variable. Column (4) shows results with training and misconduct variables winsorized at 1% level (top and bottom 0.5%). Column (5) shows results excluding January observations. T-statistics are based on standard errors clustered by country*topic and by time. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively (two-tailed).

	(1)	(2)	(3)	(4)	(5)
	<i>Misconduct</i>	<i>Misconduct</i>	$\log(Misconduct)$	<i>Misconduct</i>	<i>Misconduct</i>
	(OLS)	(OLS)	(OLS)	(OLS)	(OLS)
$In - person_small_t (thou)$	-0.516 (-1.206)	-0.481 (-1.072)	-0.416 (-1.420)	-0.615 (-1.532)	-0.534 (-1.206)
$In - person_small_{t-1} (thou)$	-0.815** (-2.412)	-0.855** (-2.504)	-0.541** (-2.165)	-0.787** (-2.362)	-0.943*** (-2.718)
$In - person_small_{t-2} (thou)$	-1.102*** (-3.407)	-1.089*** (-3.128)	-0.624** (-2.399)	-0.949*** (-3.058)	-0.783 (-1.509)
$In - person_small_{t-3} (thou)$	-0.662 (-1.519)	-0.533 (-1.231)	-0.506* (-1.721)	-0.635 (-1.564)	-0.875* (-1.796)
$In - person_big_t (thou)$	0.013 (0.739)	0.014 (0.789)	0.010 (0.875)	0.012 (0.594)	0.013 (0.762)
$In - person_big_{t-1} (thou)$	-0.020* (-1.903)	-0.021** (-2.085)	-0.016** (-2.108)	-0.028** (-2.031)	-0.025* (-1.983)
$In - person_big_{t-2} (thou)$	-0.017 (-1.417)	-0.017 (-1.411)	-0.014 (-1.620)	-0.026* (-1.736)	-0.026** (-2.010)
$In - person_big_{t-3} (thou)$	-0.003 (-0.394)	-0.003 (-0.328)	-0.003 (-0.541)	-0.005 (-0.469)	-0.003 (-0.353)
Observations	29,070	10,146	38,760	38,760	36,040
R-squared	0.096	0.078	0.445	0.307	0.463
Robustness test type	Exclude HR and Product	Include country- topics w/ both training mediums	Log	Winsor	Exclude January
Video training controls	Yes	Yes	Yes	Yes	Yes
Country-Topic FE	Yes	Yes	Yes	Yes	Yes
Time FE	-0.516	Yes	Yes	Yes	Yes

Table A7. Robustness test: Compliance training and misconduct under different levels of performance pressure

This table presents robustness test results for Table 7. Column (1) shows results excluding Employee relations/HR and Product observations. Column (2) shows result with only the country-topic matches for which both in-person training and video training was conducted during the sample period. Column (3) shows results with $\log(1+Misconduct)$ as the dependent variable. Column (4) shows results with training and misconduct variables winsorized at 1% level (top and bottom 0.5%). Column (5) shows results excluding January observations. T-statistics are based on standard errors clustered by country*topic and by time. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively (two-tailed).

	(1)			(2)		
	Production target		p-value	Production target		p-value
	Miss	Meet		Miss	Meet	
	<i>Misconduct</i>	<i>Misconduct</i>		<i>Misconduct</i>	<i>Misconduct</i>	
	(OLS)	(OLS)		(OLS)	(OLS)	
$In - person_t$ (thou)	0.021 (0.888)	0.179 (0.688)	0.393	0.022 (0.948)	0.210 (0.817)	0.474
$In - person_{t-1}$ (thou)	-0.015* (-1.990)	-0.161* (-1.859)	0.062*	-0.023** (-2.479)	-0.165** (-2.046)	0.084*
$In - person_{t-2}$ (thou)	-0.004 (-0.517)	-0.135** (-2.117)	0.633	-0.004 (-0.453)	-0.129* (-1.849)	0.081*
$In - person_{t-3}$ (thou)	0.015 (0.779)	-0.066 (-1.540)	0.444	0.016 (0.886)	-0.132** (-2.547)	0.009***
$Video_t$ (thou)	-0.006 (-0.509)	0.011 (0.352)	0.63	-0.018*** (-3.126)	0.029 (0.635)	0.296
$Video_{t-1}$ (thou)	-0.008 (-1.613)	0.010 (0.369)	0.168	-0.015 (-1.614)	-0.004 (-0.092)	0.787
$Video_{t-2}$ (thou)	-0.003 (-0.412)	-0.017** (-2.479)	0.694	0.000 (0.001)	-0.028*** (-2.922)	0.066*
$Video_{t-3}$ (thou)	-0.007 (-1.094)	0.029 (0.708)	0.306	-0.026*** (-5.047)	-0.021*** (-2.790)	0.549
Observations	6,024	5,274		2,806	1,876	
R-squared	0.139	0.070		0.105	0.084	
Robustness test type	Exclude HR and Product			Include country-topics w/ both training mediums		
Country-Topic FE	Yes	Yes		Yes	Yes	
Time FE	Yes	Yes		Yes	Yes	

Table A7. Robustness test (Continued)

	(3)			(4)		
	Production target			Production target		
	Miss	Meet	p-value	Miss	Meet	p-value
	log(<i>Misconduct</i>)	log(<i>Misconduct</i>)		<i>Misconduct</i>	<i>Misconduct</i>	
	(OLS)	(OLS)		(OLS)	(OLS)	
<i>In – person_t (thou)</i>	0.017 (1.115)	0.066 (0.460)	0.738	0.027 (0.920)	0.023 (0.172)	0.975
<i>In – person_{t-1} (thou)</i>	-0.016** (-2.189)	-0.114* (-1.859)	0.121	-0.021 (-1.559)	-0.144* (-1.973)	0.108
<i>In – person_{t-2} (thou)</i>	-0.004 (-0.647)	-0.089** (-2.077)	0.055*	-0.014 (-1.231)	-0.127** (-2.232)	0.061*
<i>In – person_{t-3} (thou)</i>	-0.001 (-0.033)	-0.043 (-1.359)	0.236	0.024 (0.572)	-0.065 (-1.519)	0.143
<i>Video_t (thou)</i>	-0.017 (-1.103)	-0.010 (-0.446)	0.759	-0.002 (-0.128)	-0.007 (-0.196)	0.910
<i>Video_{t-1} (thou)</i>	0.019 (0.765)	0.062 (1.186)	0.261	-0.001 (-0.134)	0.075 (1.297)	0.179
<i>Video_{t-2} (thou)</i>	-0.018 (-0.740)	-0.020* (-1.844)	0.967	0.003 (0.134)	-0.029* (-1.981)	0.204
<i>Video_{t-3} (thou)</i>	-0.021*** (-6.171)	0.029 (1.108)	0.029**	-0.020 (-1.074)	0.035 (1.158)	0.208
Observations	8,032	7,032		8,032	7,032	
R-squared	0.625	0.122		0.485	0.116	
Robustness test type		Log			Winsor	
Country-Topic FE	Yes	Yes		Yes	Yes	
Time FE	Yes	Yes		Yes	Yes	

Table A7. Robustness test (Continued)

	(5)		
	Production target		p-value
	Miss	Meet	
	<i>Misconduct</i>	<i>Misconduct</i>	
	(OLS)	(OLS)	
$\ln - person_t$ (thou)	0.018 (0.812)	0.429 (0.949)	0.410
$\ln - person_{t-1}$ (thou)	-0.030 (-1.542)	-0.192 (-1.681)	0.148
$\ln - person_{t-2}$ (thou)	-0.007 (-0.748)	-0.116* (-1.791)	0.085*
$\ln - person_{t-3}$ (thou)	-0.018 (-0.636)	-0.054 (-1.167)	0.273
$Video_t$ (thou)	-0.048 (-0.982)	-0.022 (-0.596)	0.782
$Video_{t-1}$ (thou)	0.115 (1.009)	0.108 (1.160)	0.325
$Video_{t-2}$ (thou)	0.069 (1.393)	-0.040 (-1.582)	0.086*
$Video_{t-3}$ (thou)	-0.062*** (-3.008)	0.042 (1.055)	0.034**
Observations	7,384	6,464	
R-squared	0.591	0.120	
Robustness test type	Exclude January		
Country-Topic FE	Yes	Yes	
Time FE	Yes	Yes	

Table A8. Robustness test: Compliance training and misconduct under different public enforcement environment

This table presents robustness test results for Table 8. Column (1) shows results excluding Employee relations/HR and Product observations. Column (2) shows result with only the country-topic matches for which both in-person training and video training was conducted during the sample period. Column (3) shows results with $\log(1+Misconduct)$ as the dependent variable. Column (4) shows results with training and misconduct variables winsorized at 1% level (top and bottom 0.5%). Column (5) shows results excluding January observations. T-statistics are based on standard errors clustered by country*topic and by time. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively (two-tailed).

	(1)			(2)		
	Enforcement		p-value	Enforcement		p-value
	Strong	Weak		Strong	Weak	
	<i>Misconduct</i>	<i>Misconduct</i>		<i>Misconduct</i>	<i>Misconduct</i>	
	(OLS)	(OLS)		(OLS)	(OLS)	
$In - person_t$ (thou)	0.246	-0.005	0.395	0.144	-0.001	0.481
	(0.851)	(-0.772)		(0.715)	(-0.083)	
$In - person_{t-1}$ (thou)	-0.130**	-0.017*	0.04**	-0.129***	-0.011	0.012**
	(-2.609)	(-1.853)		(-3.017)	(-1.233)	
$In - person_{t-2}$ (thou)	-0.120***	-0.013	0.01**	-0.135***	-0.013	0.010**
	(-3.089)	(-1.232)		(-3.022)	(-1.398)	
$In - person_{t-3}$ (thou)	-0.033	-0.004	0.652	-0.050	0.007	0.250
	(-0.520)	(-0.469)		(-1.039)	(0.761)	
$Video_t$ (thou)	-0.003	0.022	0.274	-0.004	0.013	0.564
	(-0.493)	(1.004)		(-0.501)	(0.468)	
$Video_{t-1}$ (thou)	-0.013**	0.012	0.237	-0.014	-0.015	0.970
	(-2.603)	(0.651)		(-1.625)	(-0.764)	
$Video_{t-2}$ (thou)	0.003	-0.000	0.902	0.003	0.009	0.854
	(0.268)	(-0.010)		(0.244)	(0.361)	
$Video_{t-3}$ (thou)	0.009	0.020	0.666	0.003	-0.036**	0.174
	(0.613)	(0.799)		(0.135)	(-2.613)	
Observations	14,364	14,706		5,073	5,073	
R-squared	0.054	0.106		0.070	0.095	
Robustness test type	Exclude HR and Product			Include country-topics w/ both training mediums		
Country-Topic FE	Yes	Yes		Yes	Yes	
Time FE	Yes	Yes		Yes	Yes	

Table A8. Robustness test (Continued)

	(3)			(4)		
	Enforcement		p-value	Enforcement		p-value
	Strong	Weak		Strong	Weak	
	log(<i>Misconduct</i>)	log(<i>Misconduct</i>)		<i>Misconduct</i>	<i>Misconduct</i>	
	(OLS)	(OLS)		(OLS)	(OLS)	
<i>In – person_t (thou)</i>	0.154 (0.784)	-0.005 (-0.911)	0.427	0.234 (0.821)	-0.010 (-1.050)	0.403
<i>In – person_{t-1} (thou)</i>	-0.096** (-2.446)	-0.016** (-2.101)	0.061*	-0.138** (-2.578)	-0.026** (-2.065)	0.067*
<i>In – person_{t-2} (thou)</i>	-0.084*** (-2.792)	-0.012* (-1.706)	0.027**	-0.122*** (-3.040)	-0.020 (-1.537)	0.017**
<i>In – person_{t-3} (thou)</i>	-0.033 (-0.804)	-0.001 (-0.163)	0.445	-0.034 (-0.545)	-0.003 (-0.235)	0.622
<i>Video_t (thou)</i>	-0.035 (-1.441)	0.007 (0.386)	0.275	-0.014 (-1.090)	0.009 (0.416)	0.442
<i>Video_{t-1} (thou)</i>	-0.016 (-1.390)	0.057 (1.174)	0.015**	-0.031*** (-3.374)	0.067 (1.497)	0.030**
<i>Video_{t-2} (thou)</i>	-0.015 (-0.921)	0.018 (0.890)	0.141	0.013 (0.549)	0.003 (0.108)	0.756
<i>Video_{t-3} (thou)</i>	-0.000 (-0.049)	0.000 (0.006)	0.986	0.009 (0.493)	0.006 (0.357)	0.900
Observations	19,152	19,608		19,152	19,608	
R-squared	0.432	0.454		0.284	0.319	
Robustness test type		Log			Winsor	
Country-Topic FE	Yes	Yes		Yes	Yes	
Time FE	Yes	Yes		Yes	Yes	

Table A8. Robustness test (Continued)

	(5)		
	Enforcement		
	Strong	Weak	
	Misconduct	Misconduct	
	(OLS)	(OLS)	p-value
$ln - person_t$ (thou)	0.229 (0.805)	-0.014 (-1.081)	0.406
$ln - person_{t-1}$ (thou)	-0.125** (-2.123)	-0.027 (-1.563)	0.122
$ln - person_{t-2}$ (thou)	-0.122** (-2.524)	-0.028 (-1.643)	0.085*
$ln - person_{t-3}$ (thou)	-0.059 (-0.885)	0.002 (0.175)	0.369
$Video_t$ (thou)	-0.084 (-1.394)	-0.011 (-0.268)	0.363
$Video_{t-1}$ (thou)	-0.002 (-0.067)	0.125 (1.090)	0.01**
$Video_{t-2}$ (thou)	0.030 (1.113)	0.135 (1.094)	0.401
$Video_{t-3}$ (thou)	-0.017 (-0.867)	-0.026 (-0.443)	0.915
Observations	17,808	18,232	
R-squared	0.450	0.471	
Robustness test type	Exclude January		
Country-Topic FE	Yes	Yes	
Time FE	Yes	Yes	

Table A9. Compliance training and misconduct by training medium (by topic)

This table presents OLS regression results that test the effect of compliance training on misconduct by training medium for three misconduct topics that have at least 50 cases of misconduct during the sample period: Employee relations/HR, Financial reporting/Records management, and Procurement/ Bid rigging. Variables are defined in the header of Table 3. The dependent variable in Panel A is the number of substantiated allegations. T-statistics are based on standard errors clustered by country-topic and by time. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively (two-tailed).

	(1) Employee relations/ HR <i>Misconduct</i> (OLS)	(2) Financial reporting/ Records management <i>Misconduct</i> (OLS)	(3) Procurement/ Bid rigging <i>Misconduct</i> (OLS)
$\ln - person_t$ (thou)		-0.105* (-1.708)	-0.189** (-2.013)
$\ln - person_{t-1}$ (thou)		-0.103 (-1.300)	-0.328* (-1.975)
$\ln - person_{t-2}$ (thou)		-0.115 (-1.400)	-0.379** (-2.090)
$\ln - person_{t-3}$ (thou)		-0.010 (-0.069)	-0.025 (-0.078)
$Video_t$ (thou)	-0.157 (-1.475)	0.021 (0.372)	-0.034* (-1.991)
$Video_{t-1}$ (thou)	0.079 (0.388)	-0.004 (-0.108)	0.004 (0.153)
$Video_{t-2}$ (thou)	-0.029 (-0.135)	-0.013 (-0.232)	0.028 (1.042)
$Video_{t-3}$ (thou)	-0.040 (-1.658)	0.156*** (4.772)	0.049* (1.882)
Observations	4,845	4,845	4,845
(Pseudo) R-squared	0.490	0.127	0.095
Country FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes

Table A10. Ethics training and misconduct

This table presents OLS and zero-inflated negative binomial regression results that test the effect of ethics training on misconduct. *In-person_Ethics* (*Video_Ethics*) is the number of employees who received in-person (video) training on ethics. T-statistics are based on standard errors clustered by country-topic and by time. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively (two-tailed).

	(1) <i>Misconduct</i> (OLS)	(2) <i>Misconduct</i> (ZINB)
<i>In – person_Ethics_t</i> (<i>thou</i>)	0.107 (1.437)	0.687 (0.958)
<i>In – person_Ethics_{t-1}</i> (<i>thou</i>)	-0.032 (-0.991)	-0.593 (-0.563)
<i>In – person_Ethics_{t-2}</i> (<i>thou</i>)	0.118 (1.383)	0.597 (1.245)
<i>In – person_Ethics_{t-3}</i> (<i>thou</i>)	-0.059 (-1.600)	-1.888* (-1.929)
<i>Video_Ethics_t</i> (<i>thou</i>)	0.002 (0.232)	0.064 (0.622)
<i>Video_Ethics_{t-1}</i> (<i>thou</i>)	0.004 (0.351)	0.098 (1.001)
<i>Video_Ethics_{t-2}</i> (<i>thou</i>)	0.008 (1.048)	0.103 (1.205)
<i>Video_Ethics_{t-3}</i> (<i>thou</i>)	-0.004 (-0.591)	0.088 (1.279)
Observations	38,760	38,760
(Pseudo) R-squared	0.459	0.401
Compliance Training controls	Yes	Yes
Country-Topic FE	Yes	
Time FE	Yes	
Inflator		Country-Topic, Time