Artificial Intelligence and Workplace Misconduct

Suleyman Faruk Gozen[*] Mehmet Furkan Karaca[†] Monika Tarsalewska[‡]

Abstract

This paper studies the impact of artificial intelligence (AI) adoption on workplace misconduct. Using regression analysis and quasi-natural experiment around the open-source launch of Google Brain's TensorFlow machine learning toolkit, we find that firms with higher AI intensity experience significant and persistent declines in workplace violations and penalty amounts. The effects operate primarily through productivity-enhancing complementarities and discretionary expenses increases, while labor-adjustment channels play no role. Benefits are concentrated among larger, intangible- and organizational-capital intensive firms, which highlights uneven gains from AI adoption and suggests that AI may widen compliance inequality among firms.

Keywords: Artificial Intelligence, Workplace Misconduct, Causal Inference, Firm-Level Analysis

JEL Codes: D20, G30, J53, L20, M54, O33

[*]School of Economics, University of Bristol. 12 Priory Road, Bristol BS8 1TU, United Kingdom. Email: suleyman.gozen@bristol.ac.uk.

[†]Essex Business School, University of Essex. Wivenhoe Park, Colchester CO4 3SQ, United Kingdom. Email: m.f.karaca@essex.ac.uk.

[‡]University of Exeter Business School. Streatham Court, Streatham Campus, Rennes Drive, Exeter EX4 4PU, United Kingdom. Email: m.tarsalewska@exeter.ac.uk.

1 Introduction

The rapid diffusion of artificial intelligence (AI) is transforming firm operations through automating tasks, monitoring performance, and augmenting decision-making. Brynjolfsson et al. (2025) suggest that we need to deepen our understanding of the economic implications of AI on society. Although a growing literature has examined the effects of AI on productivity, innovation, and firm value, much less is known about how AI adoption influences employees.

AI adoption raises important questions about employee safety and labor law compliance. On the one hand, large investment in AI could crowd out traditional investments in worker training, health, and safety programs. When companies are driven by AI-productivity gains, they may neglect workplace safety measures and even push legal boundaries in employment practices. On the other hand, AI technologies, implemented with complementary strategies, might enhance workplace safety and help firms comply with labor standards.

In this paper, we aim to understand the consequences of AI adoption on labor law violations that we refer to as workplace misconduct that includes safety and employment related labor violations. Labor law violations impose substantial financial and reputational costs on firms and generate significant social harm. Understanding whether and how AI affects workplace misconduct provides new insights into the broader economic and societal consequences of technological change.

To address these questions, we adopt a two-step empirical strategy. We first document baseline correlations between AI adoption and workplace misconduct using panel regressions with fixed effects. Then, following the quasi-experimental insight from Rock (2019), we turn to a causal identification strategy that takes advantage of the open-source launch of Google Brain's TensorFlow in November 2015 as a laboratory to study the impact of AI adoption, exploiting this plausibly exoge-

nous shock within a difference-in-differences framework. The dynamic difference-in-differences framework allows us to trace the dynamic impact of AI on workplace misconduct, validate the parallel trends assumption, and distinguish genuine treatment effects from spurious correlations.

Our analysis uses data on violations and the ensuing penalties issued by federal regulatory agencies and the United States Department of Justice from Violation Tracker. Firm-level financial and governance information is drawn from Compustat, ExecuComp, BoardEx, and CRSP. AI adoption is measured following Babina et al. (2024), using the share of employees with AI-related skills to capture the extent to which firms integrate AI into their workforce. Our final sample consists of 7,906 firm-year observations across 1,107 distinct firms from 2010 to 2019.

We obtain four sets of results. First, baseline estimates indicate that firms with greater AI intensity experience significantly fewer safety and employment related violations and lower penalty amounts. These associations remain robust after controlling for industry-year, and firm fixed effects. Second, our causal analysis shows that firms with higher pre-treatment AI intensity experienced sharper declines in workplace misconduct after the 2015 release of TensorFlow. We estimate that a one standard deviation increase in AI share is associated with an additional decline of around 3% in workplace violations and about 23% in penalty amounts. Dynamic DiD results confirm the parallel trends assumption and demonstrate that the effects accumulate over time, consistent with the gradual integration of AI into compliance processes. Therefore, we find that firms with greater AI adoption are more effective in reducing workplace misconduct in the years following the TensorFlow shock.

Third, our mechanism analysis highlights the role of productivity such that firms with higher labor productivity experience sharper declines in workplace misconduct due to higher AI adoption. In addition, firms that increase discretionary expenditures exhibit greater declines, suggesting that

complementary investments such as training, monitoring, and employee support allow AI adoption to reduce misconduct. Importantly, we find no evidence that the effect of AI on workplace misconduct operates through labor-adjustment channels, such as layoffs.

Fourth, heterogeneity analysis shows that the reductions in workplace misconduct are concentrated among larger firms and those with higher intangible and organizational capital. These firms benefit from economies of scale, richer data environments, and stronger managerial capacity to integrate AI into oversight structures, while smaller or less intangible-intensive firms derive fewer compliance benefits.

We conduct an extensive set of robustness checks to validate our findings. First, we re-estimate our models across a broad range of alternative offense types beyond safety- and employment-related violations, and we find no systematic effects. Second, we show that our results are not driven by spurious timing by implementing a placebo test, which yields null effects. Third, we estimate specifications that incorporate additional corporate governance control variables as well as industry- and state-specific linear time trends to account for potential differential compliance trajectories across sectors and regions, and the results remain unchanged. Fourth, we replicate our benchmark analysis using an alternative narrow measure of AI skills, and the results remain robust. Fifth, while our benchmark sample excludes firms from Finance, Utilities, and Other highly regulated industries due to their distinct regulatory environments, we re-estimate our baseline specifications by including these sectors and find that the estimated effects remain robust. Sixth, we explicitly address selection concerns by implementing propensity score matching, comparing AI-adopting firms to observationally similar non-adopters based on pre-treatment firm characteristics; the estimated effects remain economically and statistically similar. Lastly, we show that our results are not driven by broader technological upgrading by controlling for other firm-level technologies,

including robot adoption, IT intensity, data analytics, and data infrastructure, confirming that AI operates through a distinct channel. Hence, we demonstrate that the documented effects of AI adoption on workplace misconduct are highly robust and not an artifact of modeling assumptions, sample construction, or confounding technological trends.

This paper makes the following contributions to the previous literature. First, we contribute to the growing literature on the determinants of employee welfare and workplace misconduct. Prior research finds that firms experience more workplace misconduct when they are financially constrained and not publicly listed (Cohn and Wardlaw (2016) and Liang et al. (2023)), awarding more equity incentives to CEOs (Chircop et al. (2025)), facing pressure to meet or beat quarterly earnings expectations (Caskey and Ozel (2017) and Raghunandan (2021)), subject to greater information frictions (Raghunandan and Ruchti (2024)). Contrary to expectations, firms' claims of social responsibility do not mitigate workplace misconduct (Raghunandan and Rajgopal, 2024). Firms experience less workplace misconduct when they conduct direct monitoring from headquarters (Heese and Pérez-Cavazos (2020)), when rank-and-file employees hold stock options (Chen et al. (2025)), implement mandated reporting of violations in place (Christensen et al. (2017)), voluntarily issue corporate accountability reporting (Christensen (2016)), and have higher institutional ownership (Li and Raghunandan (2021)). Our paper is closest to Heese and Pacelli (2025). They show that broader Information Technology (IT) adoption such as Enterprise Resource Planning (ERP) systems and digital recordkeeping reduces the incidence of corporate misconduct by enhancing transparency and auditability. While their focus is on general purpose IT systems that improve internal control environments and monitoring capabilities, we study AI adoption, which represents a different technological transformation. AI systems do not merely record or transmit information; they can augment employees' skills, and they can respond to compliance risks in real

time. Importantly, we show that our results are not driven by broader technological upgrading. In robustness analyses, we explicitly control for firms' exposure to other technologies, including the share of robot adoption, IT intensity, data analytics, and data infrastructure. The estimated effects of AI adoption remain stable, indicating that AI operates through a distinct channel rather than proxying for general automation, digitization, or data capacity. Thus, our paper moves beyond prior work on IT that focuses on automation and digitization.

Our study also differs from Ciconte et al. (2024a) which focus on the recent regulatory oversight and find that firms operating under stricter AI regulations have fewer employment related offenses. We investigate the internal adoption of AI technologies as an organizational decision by exploiting a quasi-natural experiment (the 2015 TensorFlow release) and provide causal evidence on how firm-level internal AI integration, independent of regulatory shocks, affects workplace misconduct. Ciconte et al. (2024a) complement our study by showing that the recent introduction of AI regulatory oversight further improves the beneficial effects of AI adoption on compliance with labor law.

Second, we contribute to the emerging literature on the relationship between AI and firm dynamics. Recent studies show that AI is linked to higher firm value (Rock (2019), Chen and Srinivasan (2024) and Babina et al. (2024)), affects financial disclosure (Ashraf (2025), Ding et al. (2020), Fedyk et al. (2022) and Akyildirim et al. (2025)), and greater innovation (Cockburn et al. (2018)), while AI laws hinder innovation (Ciconte et al. (2024b)). We contribute to this literature by investigating potential channels that can impact our main results on the relationship between AI and workplace misconduct. We also uncover important heterogeneity that the compliance benefits of AI are concentrated among large, intangible-, and organizational-capital intensive firms, highlighting uneven gains from technological change.

Our findings connect to recent calls to understand how emerging technologies reshape organizational culture and internal governance. Grennan and Li (2023) highlight that advances in AI-enabled "smart" technologies may take on managerial functions, alter employee expectations, and influence how cultural norms are reinforced. They explicitly raise the question of whether AI-based decision aids can improve corporate outcomes or instead affect employees' motivation and trust. Our evidence provides a specific mechanism relevant to this question: by enhancing monitoring capacity and complementing firms' organizational capital, AI adoption reduces workplace misconduct without inducing labor displacement. This suggests that AI can operate as a governance-enhancing technology, strengthening internal norms and compliance, which is precisely the type of mechanism Grennan and Li (2023) argue warrants empirical investigation.

Our findings have important policy implications. We show that AI technologies can act as a governance enhancing mechanism, not merely an efficiency tool. Our results suggest that AI adoption generates positive externalities by curbing workplace misconduct, i.e., safety and employment related violations. At the same time, the heterogeneous effects imply that these benefits are unevenly distributed, raising concerns that smaller firms may lag behind on compliance improvements. Policymakers may therefore consider ways to support the diffusion of AI adoption across a broader range of firms to ensure more equitable improvements in workplace standards.

We organize the paper as follows. Section 2 reviews the related literature and develops our hypotheses. Section 3 describes the data and variable construction. Section 4 presents baseline descriptive and regression evidence. Section 5 establishes causality and presents the main results. Section 6 examines heterogeneity and conducts robustness checks. Section 7 concludes with key findings and policy implications.

2 Literature Review and Hypotheses Development

This section reviews the relevant literature on artificial intelligence, labor markets, and workplace misconduct, and develops the hypotheses that guide our empirical analysis.

2.1 Artificial Intelligence and Labor Markets

The early literature on labor economics viewed the arrival of new technology as a threat to employment. Yet, Autor (2015) argues that while automation can be a substitute for labor, it is also a complement to labor as it creates demand for new skills. He also suggests that while machines can replace humans in performing routine, and codifiable tasks; this allows humans to excel in problem-solving skills, adaptability, and creativity.

Recent literature on AI technology highlights both disruptive and complementary effects on labor markets. Studies of online freelancing platforms show immediate substitution effects (Hui et al. (2024) and Demirci et al. (2025)), while firm-level evidence suggests that generative AI increases productivity by 15% in customer service roles, with the largest benefits accruing to less-experienced and lower-skilled workers (Brynjolfsson et al. (2025)).

Generative AI affects many occupations, with as many as 80% of U.S. jobs containing tasks that could be affected (Eloundou et al. (2024), Felten et al. (2021) and Tomlinson et al. (2025)). AI is more likely to augment than fully automate most roles (Gmyrek et al. (2023)). Yet, Brynjolfsson et al. (2025) point out significant displacement of young workers in AI-exposed occupations. At the same time, analyses of large-scale AI usage data indicate that most adoption involves collaborative use cases such as writing, coding, and information gathering where AI assists rather than replaces human labor (Handa et al. (2025)).

Brynjolfsson et al. (2025) show that employment declines occur in jobs where AI is more likely to automate than augment human labor. Thus, AI adaptation, rather than eliminating employment, should reshape workplaces by augmenting human skills. Early evidence on AI adoption by Babina et al. (2024) show that companies investing in AI have higher sales, employment, and market valuations relative to peers. These changes occur mainly not due to cutting labor costs or workforce replacements but due to product innovation. Firms using AI were able to introduce new products, hold patents and trademarks, and achieve better product quality that increased demand.

AI can therefore enhance employee productivity, improve decision-making quality, and foster innovation. These findings imply that AI's labor-market effects are multi-dimensional and combine substitution, augmentation, and coordination dynamics.

2.2 AI Adoption and Workplace Misconduct

Prior research links a greater incidence of workplace misconduct to financial pressures, managerial incentives, information frictions, and weak monitoring (Cohn and Wardlaw (2016), Caskey and Ozel (2017), Raghunandan (2021), Chircop et al. (2025), Raghunandan and Ruchti (2024) and Heese and Pérez-Cavazos (2020)). In terms of technological adaptation, Heese and Pacelli (2025) show that improvements in digital tools reduce corporate misconduct by enhancing transparency. Similarly, Ciconte et al. (2024a) find that regulatory oversight of AI technology reduces employment-related misconduct, consistent with other studies emphasizing the role of improved monitoring.

AI is transforming how firms monitor employees, allocate resources, and enforce compliance. Yet, its impact on workplace misconduct remains theoretically ambiguous. On the one hand, AI

may strengthen governance, transparency, and ethical oversight. On the other hand, it may introduce new risks by eroding human judgment, amplifying managerial pressure, or reducing investment in worker welfare.

AI adoption can enhance labor law compliance and reduce workplace misconduct through complementarity and organizational learning, as well as monitoring and deterrence. AI systems complement human decision-making by improving data-driven planning, resource allocation, and process control. They can also augment employees' skills. Gihleb et al. (2022) show that, for U.S. manufacturing firms, an increase in the use of industrial robots is associated with lower employee injury rates. In the case of AI adoption, Babina et al. (2024) find that firms investing in product innovation experience higher demand, sales, and employment. Therefore, when firms combine AI adoption with investments in employee training and safety practices, e.g., teaching workers how to operate equipment more safely, we can expect fewer workplace violations. For example, AI systems can improve compliance by monitoring employees for risks (e.g., computer vision that detects unsafe behaviors likely to cause accidents).

AI adoption may also improve compliance through organizational learning. For instance, AI-driven efficiency can accelerate task completion, and reduce overtime pressure on employees. Eloundou et al. (2024) suggest that the majority of U.S. employees could have at least 10% of their work tasks affected by large language models (LLMs), which significantly increase task speed. Thus, firms that combine AI adoption with complementary investments, such as employee training, safety programs, and monitoring systems are more likely to benefit more from AI in improving compliance effectively.

AI systems can also enhance labor law compliance through monitoring and deterrence. AI is not only a production technology but also a governance technology that alters the costs, benefits,

and observability of workplace misconduct. Becker (1968) argues that individuals comply with rules when the expected cost of misconduct exceeds the expected benefit. Since enhanced monitoring reduces corporate misconduct (Heese and Pérez-Cavazos (2020)), AI adoption can further lower violations by increasing detection probability and deterrence. We therefore expect that AI adoption may decrease the incidence and severity of labor and safety violations.

Conversely, AI adoption may also lower labor law compliance and increase workplace misconduct through resource reallocation, neglect of human oversight, and managerial short-termism. When firms rely excessively on technology to monitor or replace workers, they may cut training budgets and underinvest in safety programs. This crowding-out effect can degrade working conditions and inadvertently increase the likelihood of violations. For instance, the United States Senate Committee on Health, Education, Labor, and Pensions (HELP) (2024) reports that some safety law violations stemmed from automation and productivity monitoring, highlighting that when mismanaged, AI adoption may lead to underinvestment in employee safety and greater legal non-compliance.

AI-based monitoring systems may also create new forms of misconduct risk by generating false positives or enforcing rigid performance metrics. Without adequate human oversight, algorithmic systems can misclassify employee behavior, incentivize performance gaming, or exacerbate workplace inequities, thereby eroding ethical culture. The U.S. Department of Labor's (DOL) Wage and Hour Division (United States Department of Labor, Wage and Hour Division (2024)) reports that employers using AI-based monitoring of hours and performance are at greater risk of violating wage and hour laws. In 2024, the DOL issued guidance noting that automated scheduling, time-tracking, and productivity-scoring tools can lead to systemic violations of the Fair Labor Standards Act (FLSA) if not supervised by humans. AI surveillance can also erode employee morale

and mental health (Gihleb et al. (2022)), leading indirectly to more safety incidents.

AI investments might therefore undermine organizational complementarities by diverting resources from human capital development and excessive performance monitoring, ultimately reducing compliance with labor laws. We therefore conjecture that, under certain conditions, AI adoption may increase the incidence and severity of labor and safety violations.

Given these competing predictions, our empirical analysis will assess which effect prevails.

3 Sample and Variable Measurement

This section describes the data sources, sample construction, and variable definitions underlying our empirical analysis, including the measures of AI adoption, workplace misconduct, firm characteristics and the fixed effects structure used in the regressions.

3.1 Sample Construction

Our sample is constructed mainly from Compustat, a dataset providing financial information for U.S. public companies. We start by collecting firm-level financial information from Compustat covering the years 2010 to 2019. The sample period is selected to accommodate the quasi-experimental empirical research design. We apply standard screens following Chircop et al. (2025) to exclude firms with Standard Industrial Classification (SIC) codes between 4900 and 4999 (utilities), between 6000 and 6999 (financial institutions), and above 9000 (public administration and nonclassifiable establishments).[1] For workplace misconduct, we use data from Violation Tracker and focus on safety and employment related offense types. We merge these data with our sample

[1] We provide a robustness check in Section 6, where we include these industries in the analysis.

and construct necessary control variables for our baseline regressions. The resulting sample in our causal inference regression comprises 7,906 firm-year observations across 1,107 distinct firms. All monetary values are normalized to 1990 U.S. dollars using the Consumer Price Index (CPI). All variables are winsorized at the 1st and 99th percentiles within each year. The sample selection process is discussed in Table A.1.

3.2 Variable Measurement

3.2.1 Measurement of Artificial Intelligence

We rely on the measures constructed by Babina et al. (2024) to quantify the proportion of AI-related employees at the firm level. Their study develops a novel proxy for corporate AI investment by analyzing the extent of hiring for AI-specific skills using detailed information from employee résumés and job postings. This measure reflects both the existing pool of AI talent within firms and their ongoing demand for AI expertise. Furthermore, they distinguish between "narrow" and "broad" definitions of AI skills. The narrow measure excludes skills with limited direct relevance to AI. In our analysis, we use the broad definition, as it captures a wider range of AI-related expertise and better reflects the diverse applications of AI across firms and industries. As a robustness check, we also re-estimate our benchmark analysis using the narrow definition and find that the results remain consistent.

3.2.2 Measurement of Workplace Misconduct

To measure workplace misconduct, we draw data from Violation Tracker, a database maintained by the Corporate Research Project of Good Jobs First. This resource compiles a wide range of viola-

tions segregated into several categories related to competition, consumer protection, employment, environmental, safety related violations, among others.

Violation Tracker reports only cases with fines exceeding $5,000. For our analysis, we keep only safety- and employment-related offenses, as these categories are most directly connected to the channels through which AI adoption can affect employees. If no safety and employment related offenses are reported for a firm in a given year, we replace the missing value with zero. This approach ensures that the absence of reported violations is treated as a meaningful outcome, rather than being removed from the analysis. Thus, we preserve the full panel structure. Figure B.1 shows histograms of workplace misconduct variables (number of violations and penalty amounts), while Figure B.2 presents their averages across Fama–French industries. Both figures reveal substantial variation and heterogeneity in terms of workplace misconduct variables across firms and industries in our sample.

3.2.3 Control Variables

We incorporate a set of time-varying firm characteristics that prior studies have linked to misconduct (Cohn and Wardlaw (2016) and Chircop et al. (2025)). Specifically, for our main tests we include measures of *Leverage, CashFlow, DividendPayout, Size, Sales, Employees, LaborIntensity, AssetTurnover, Market-to-Book, Tangibility, Capex, R&D Intensity*, and *FirmAge*. *Leverage* is defined as the book value of debt divided by the book value of assets. *CashFlow* is operating cash flows scaled by total assets. *DividendPayout* is the ratio of cash dividends to common shareholders over lagged book assets. *FirmSize* is the natural logarithm of market capitalization. *Sales* is annual growth in firm sales. *Employees* is the natural logarithm of the number of employees. *LaborIntensity* is employees per unit of total assets. *AssetTurnover* is sales over total assets. *Market-to-Book*

is the market-to-book value of assets. *Tangibility* is the ratio of tangible fixed assets to total assets. *Capex* is capital expenditures on long-term tangible assets (excluding acquisitions), scaled by lagged assets. *R&D Intensity* is R&D expenditures scaled by total assets. *FirmAge* is the difference between the current year and the firm's founding year.[2]

For additional tests, we incorporate measures of intangible capital following Peters and Taylor (2017), which constructs firm-level stocks using a perpetual inventory approach. In addition to total intangible capital, we specifically consider organizational capital, derived from a portion of SG&A expenses that reflect investments in managerial structures and routines. Both measures are matched to our main sample.

For additional tests to our main sample we add data on corporate governance from BoardEx to calculate the total number of directors at the firm-year level (*Number of Directors*) and the ratio of non-executive directors on each company's board (*Non-Executive Director Ratio*).

We also add data on executive compensation and incentives from ExecuComp and CRSP. We construct measures of risk-taking incentives (*Vega*), stock price sensitivity (*Delta*), and *FirmRelatedWealth* following Coles et al. (2006, 2013) and Chircop et al. (2025). Using the dataset provided by Coles et al. (2006)[3], we compute the median values of the executives measures at the firm–year level. In that respect, *Vega* captures the change in the median value of executives' wealth resulting from a 0.01 increase in the volatility of firm stock returns. *Delta* measures the dollar change in the median value of executives' wealth associated with a 1 percent change in the firm's stock price. *FirmRelatedWealth* reflects the median value of executives' equity and option holdings. We

[2]We obtain firms' founding year from Jay Ritter's website. However, it is not available for all firms in Compustat. We use the first year reported in CRSP as the founding year when Jay Ritter's data is not available. Next, we use the first year appeared in Compustat if both of the former are not available. Lastly, we use the earliest year among the first year appeared in Compustat, the first year reported in CRSP, and the founding year from Jay Ritter's data to construct founding year for firms when all three sources are available.

[3]The dataset is available here.

also include CEO-level variables: *CEOage* (in years) and *CEOtenure* (years since appointment) to capture managerial characteristics.

To account for unobserved heterogeneity, we include firm fixed effects to control for time-invariant firm characteristics, industry-by-year fixed effects to capture industry-specific time trends in misconduct. This specification yields a within-firm analysis of the relationship between AI adoption and workplace misconduct, thereby mitigating concerns about omitted variable bias.

Table A.2 summarizes the variables used in the analysis, along with definitions, unit of observations, and data sources.

3.3 Sample Distribution

Table 1 summarizes the distribution of firm-year observations across years (Panel A) and industries (Panel B). The yearly number of observations remains relatively stable. Panel B reports the breakdown across the 12 Fama-French industries. The largest shares of the sample are drawn from High Tech industries including computers, software, and electronic equipment (32.3%), Healthcare including medical equipment and drugs (21.4%), and Manufacturing (14.3%). In contrast, industries such as Oil, Gas, and Coal (0.8%) and Telephone and Television (1.6%) are least represented.

Table 2 presents descriptive statistics for the variables used in our analysis. Our main outcomes capture both the frequency of workplace misconduct and the associated penalty amounts. The raw distributions are highly skewed: while the mean number of violations is around 0.28 per firm-year, the median is zero, indicating that most firms do not report any violations in a given year, but a small subset of firms incur very large counts or penalties. Similarly, the average penalty exceeds $6,000, yet the 75th percentile is zero, showing that substantial fines are concentrated among a

limited group of firms. This motivates the use of natural log transformations (hereafter, denoted by *log*) to mitigate skewness and improve comparability across firms.

The table also summarizes the control variables. Firms exhibit wide heterogeneity in financial characteristics such as leverage, sales growth, and market-to-book ratios, with large standard deviations pointing to heavy-tailed distributions. Corporate governance variables (e.g., *Vega*, *Delta*, *Firm-Related Wealth*) also display strong dispersion, consistent with variation in managerial incentives across firms. Intangible capital measures reveal that firms differ substantially in their total intangible and organizational capital stocks, while AI adoption, as measured by the share of AI workers, remains small on average but with meaningful variation across firms.

Table A.3 reports pairwise correlations among the variables used in our analysis. As expected, our two key outcome measures (the natural logarithm of one plus the number of violations, and the natural logarithm of one plus the penalty amounts) are strongly correlated, reflecting that firms with more violations also face larger fines. Firm size and employment measures are positively correlated with misconduct outcomes, consistent with larger firms being more visible to regulators and therefore more likely to be penalized. In contrast, labor productivity is negatively correlated with misconduct, suggesting that more efficient firms tend to have fewer violations. Governance variables such as the number of directors, the non-executive director ratio, *Delta*, and *Vega* show modest positive correlations with misconduct. Overall, the correlations are intuitive and reinforce the importance of controlling for key firm and corporate governance characteristics in our regression analysis.

4 Baseline Results

This section provides baseline evidence on the relationship between AI adoption and workplace misconduct using descriptive plots and regression analysis.

Figure B.3 provides initial visual evidence on the relationship between AI adoption and workplace misconduct. Panel (a) plots the share of AI workers against the natural logarithm of one plus the number of violations, while Panel (b) plots the same against the natural logarithm of one plus the penalty amounts. In both cases, the fitted lines reveal a negative association: firms with a higher proportion of AI-related employees tend to experience fewer violations and lower penalty amounts.

These binscatter plots serve two purposes. First, they offer a descriptive, non-parametric perspective that illustrates the raw correlation between AI intensity and misconduct outcomes, independent of the controls and fixed effects used in the regression framework. Second, they provide suggestive evidence in support of our hypothesis: that AI adoption is associated with improvements in compliance and reductions in workplace misconduct.

To move beyond descriptive patterns, we estimate panel regressions of the following form:

$$Y_{it} = \beta \, \text{AIShare}_{i,t-1} + \gamma' X_{i,t-1} + \lambda_{j,t} + \alpha_i + \varepsilon_{it}, \tag{1}$$

where Y_{it} denotes workplace misconduct outcomes for firm i in year t, measured alternatively as the natural logarithm of one plus the number of violations or the natural logarithm of one plus the penalty amounts. The key explanatory variable $\text{AIShare}_{i,t-1}$ is the lagged share of AI workers. $X_{i,t-1}$ represents a rich set of controls capturing firm characteristics. Specifically, we account

for a lag of *Leverage*, *CashFlow*, *DividendPayout*, *FirmSize*, *Sales*, *Employees*, *LaborIntensity*, *AssetTurnover*, *Market-to-Book*, *Tangibility*, *Capex*, *R&D Intensity*, and *FirmAge*.[4] The specification further includes industry-year fixed effects $\lambda_{j,t}$, and firm fixed effects α_i. Standard errors are clustered at the firm level.

Table 3 reports baseline panel regression estimates of the relationship between AI adoption, measured by the lagged share of AI workers, and workplace misconduct outcomes. Across all specifications, the coefficients on AI adoption are negative and statistically significant for the number of violations. The magnitudes are economically meaningful, indicating that greater AI intensity is associated with substantial reductions in the frequency of workplace misconduct. These effects remain robust as fixed effects are introduced, including industry-year, and firm fixed effects. This evidence suggests that the documented relationship is not merely driven by cross-sectional variation across firms or industries, but also reflects within-firm changes in AI adoption over time.

Nevertheless, the baseline regression results should be interpreted with caution. The estimates may still be influenced by endogeneity concerns, such as reverse causality, where firms with weaker compliance records adopt AI technologies as a corrective strategy, or omitted variables, for example, unobserved improvements in corporate governance that are jointly associated with both AI adoption and reductions in misconduct. Therefore, in the next section, we address potential endogeneity concerns using a causal inference framework.

[4]Because the inclusion of corporate governance variables reduces the number of observations and degrees of freedom in our regressions, we incorporate them only in the robustness analysis presented in Section 6. In this section, we include additional control variables capturing corporate governance characteristics, such as firm-level risk-taking incentives (*Vega*), stock price sensitivity (*Delta*), firm-related wealth (*FirmRelatedWealth*), CEO age (*CEOage*) and tenure (*CEOtenure*), as well as board-level governance measures, including the number of directors (*Number of Directors*) and the ratio of non-executive directors (*Non-Executive Director Ratio*).

5 Addressing Endogeneity

This section presents our efforts to establish causality on the effects of AI adoption on workplace misconduct by exploiting the 2015 release of Google TensorFlow as a plausibly exogenous shock to the effectiveness of AI technologies, and by implementing a standard and dynamic difference-in-differences research design. This approach allows us to move beyond correlations and provide more credible evidence on the effect of AI adoption on workplace misconduct.

5.1 Identification Strategy

To estimate the effect of firm-level AI adoption on workplace misconduct, we exploit the 2015 release of Google TensorFlow, a widely adopted open-source platform that substantially lowered the costs and technical barriers to deploying machine learning applications. This technological breakthrough enabled firms to leverage their existing AI workforce more effectively. Anecdotal evidence suggests that post 2015 there was a sharp increase in AI skills with 190% rise of reported AI skills on LinkedIn profiles from 2015 to 2017.[5]

The 2015 release of Google TensorFlow represents a pivotal moment in the diffusion of artificial intelligence technologies. As highlighted by Rock (2019), TensorFlow substantially lowered both the financial and technical barriers to entry by providing firms with an open-source, general-purpose framework for building and deploying machine learning applications. Unlike proprietary AI tools that were costly and specialized, TensorFlow offered a flexible and scalable platform supported by extensive documentation, community contributions, and integration with existing programming environments. This democratization of AI technologies meant that firms did not need

[5]More details can be found here.

to invest heavily in proprietary infrastructure; rather, they could build on their existing capabilities and, crucially, increase the effectiveness of their AI workforce by making their skills more productive. In this sense, the release of TensorFlow acted as a plausibly exogenous shock that raised the marginal returns to AI-related human capital and amplified the value of prior AI investments. This makes it a particularly well-suited setting for our study, as it provides a source of variation in AI effectiveness that is external to firms' underlying compliance environments, allowing us to better isolate the impact of AI adoption on workplace misconduct from broader trends in performance, governance, or regulation.

A critical step in validating our identification strategy is to confirm that the 2015 release of Google TensorFlow indeed altered firms' AI labor composition. If our difference-in-differences (DiD) design correctly captures this technological discontinuity, we should observe a statistically significant increase in the share of AI-related workers in firms that were more exposed to AI technologies prior to 2015. Conversely, firms with low pre-2015 AI exposure should exhibit little to no change in their AI workforce shares.

Table 4 presents the results of this identification check. The dependent variable is the firm-level share of AI workers, while the key regressor is the interaction between the post-2015 indicator and the pre-treatment firm-level share of AI workers in 2014. Across all specifications, the coefficient on the interaction term is positive, statistically significant, and economically meaningful. The estimates range from 0.255 to 0.416, significant at the 5% or 1% level, implying that firms with higher initial exposure to AI saw substantially greater growth in their AI workforce after the TensorFlow release.

Having established this first validation test, we now capture whether workplace misconduct changed after this exogenous shock. We use the pre-treatment firm-level share of AI workers in

2014 as a proxy for ex-ante AI intensity. Our empirical strategy follows a difference-in-differences (DiD) design, estimating the following specification:

$$Y_{it} = \beta_1 \text{Post}_t \times \text{Share of AI Workers}_{i,2014} + X'_{it-1}\gamma + \lambda_{j,t} + \alpha_i + \varepsilon_{it} \tag{2}$$

where Y_{it} denotes firm i's outcome in year t, measured by either the natural logarithm of one plus the number of violations or the natural logarithm of one plus the penalty amounts. The indicator variable Post_t equals one for years following the release of TensorFlow (i.e., $t \geq 2015$), and zero otherwise. $\text{Share of AI Workers}_{i,2014}$ represents the firm's share of AI-related workers in 2014 and proxies for the firm's potential to benefit from TensorFlow's release. The interaction term $\text{Post}_t \times \text{Share of AI Workers}_{i,2014}$ identifies the average treatment effect of increased AI effectiveness on workplace misconduct.

The vector X_{it-1} comprises a set of controls capturing firm characteristics as in equation (1). The specification further includes industry-year fixed effects $\lambda_{j,t}$, and firm fixed effects α_i.[6] Standard errors are two-way clustered at the firm and year level.

We further estimate a dynamic difference-in-differences specification to examine the evolution of treatment effects over time:

$$Y_{it} = \sum_{\tau \neq -1} \beta_\tau \left(\mathbb{1}\{t = \tau\} \times \text{Share of AI Workers}_{i,2014} \right) + X'_{it-1}\gamma + \lambda_{j,t} + \alpha_i + \varepsilon_{it} \tag{3}$$

where the dependent variable Y_{it} represents firm i's outcome in year t, measured by either the

[6] In the final sample, firms with state information remain in the same state throughout the sample period; therefore, including state fixed effects would be subsumed by firm fixed effects. Approximately 3.5% of firms in the final sample lack state information, and we exclude them from the baseline regression analysis since we cannot verify whether their state location changes over time. In Section 6, where we apply alternative empirical specifications (e.g., only firm and year fixed effects), these firms are included into the analysis.

natural logarithm of one plus the number of violations or the natural logarithm of one plus the penalty amounts.

The variable Share of AI Workers$_{i,2014}$ denotes the firm's share of AI-related workers in 2014 and serves as a continuous proxy for the firm's ex-ante exposure to the 2015 release of Google TensorFlow. This continuous treatment measure provides cross-sectional variation in firms' potential to benefit from the technological advancement. The terms $\mathbb{1}\{t = \tau\}$ are event-time indicators, where $\tau = 0$ denotes the year of the TensorFlow release (2015), and $\tau = -1$ is omitted and serves as the reference period. We employ the same set of control variables described for equation (2).

5.2 Identification Assumptions

Our empirical strategy relies on the assumption that the 2015 release of Google TensorFlow represents a plausibly exogenous technological shock that differentially affected firms according to their pre-existing AI intensity. Specifically, we assume that, absent the release of TensorFlow, workplace misconduct outcomes for firms with higher pre-2015 shares of AI workers would have followed parallel trends to those of firms with lower shares of AI workers. Under this parallel trends assumption, any post-2015 divergence in misconduct outcomes between high- and low-AI-intensity firms can be attributed to the enhanced effectiveness of AI adoption induced by TensorFlow. A second key assumption is that the pre-treatment share of AI workers measured in 2014 is predetermined and not itself influenced by anticipation of TensorFlow's release. This is reasonable given the unexpected nature of TensorFlow's open-source release and the fact that firms could not have tailored their AI workforce in 2014 based on knowledge of this forthcoming event. Third, we assume that no other contemporaneous shocks in 2015 differentially affected firms based on

their pre-treatment AI intensity. To mitigate this concern, we control for a comprehensive set of firm fundamentals, governance variables, as well as industry-year, and firm fixed effects. These controls absorb time-varying confounders at multiple levels and help ensure that the estimated effect captures the differential impact of TensorFlow rather than broader economic or regulatory shifts. Finally, we assume that the treatment effect operates through the mechanism of increased AI effectiveness, i.e. TensorFlow's role in enhancing the effectiveness of AI-related human capital, rather than through unrelated channels. This assumption aligns with the interpretation in Rock (2019) that TensorFlow was a general-purpose platform that democratized AI adoption by reducing implementation costs and raising the marginal returns to firms' existing AI workforce.

5.3 Standard and Dynamic Difference-in-Differences Estimates

Table 5 presents the difference-in-differences (DiD) estimates exploiting the 2015 release of Google TensorFlow as an exogenous shock to the effectiveness of AI adoption. The results indicate that firms with higher pre-treatment AI intensity experienced larger reductions in workplace violations and penalty amounts in the post-2015 period compared with firms with little or no AI exposure. More specifically, we estimate that, given the standard deviation of our running variable (the firm-level share of AI workers in 2014) is approximately 0.0045, a one standard deviation increase in AI share is associated with an additional decline of around 3% in workplace violations and about 23% in penalty amounts, relative to firms with little or no AI exposure. The mean number of violations in our sample is around 0.28; therefore, our estimates suggest a 10% reduction in the number of violations relative to its mean value in our sample. Although this reduction may appear modest in magnitude, the social and economic implications are substantial. According to the International

Labour Organization (2025) study[7], in 2017, it is estimated that work-related illnesses and injuries represent 3.94% of global GDP per year, equivalent to about $2.99 trillion annually. Assuming a linear relation between the number of violations and the cost of workplace misconduct, if all firms decrease the number of violations by 10%, then such a decrease would translate to a decrease of $299 billion in the economic cost related to labor violations per year.

Importantly, statistical significance emerges primarily in the specifications that incorporate firm fixed effects. This pattern suggests that once unobserved, time-invariant firm characteristics (e.g. persistent differences in governance quality, compliance culture, or monitoring capacity) are accounted for, the relationship between AI adoption and workplace misconduct becomes more precisely estimated. Without controlling for such heterogeneity, cross-sectional noise may obscure the effect of AI, resulting in less precise estimates. Thus, the significance observed in the firm fixed effects specification provides stronger evidence that the post-2015 decline in misconduct is driven by within-firm changes in AI adoption rather than compositional shifts across firms.

Figure 1 presents the dynamic difference-in-differences estimates. The purpose of this analysis is twofold. First, it provides a visual and statistical assessment of the parallel trends assumption, a key identifying condition in the difference-in-differences design. Specifically, if firms with higher and lower pre-treatment AI intensity were already on diverging trajectories prior to 2015, the causal interpretation of the estimates would be undermined. Second, the dynamic difference-in-differences framework allows us to trace the evolution of treatment effects over time, rather than focusing solely on average pre- versus post-treatment differences.

Figure 1a plots the coefficients for the natural logarithm of one plus the number of violations, and Figure 1b plots the corresponding coefficients for the natural logarithm of one plus the penalty

[7]More details can be found here.

amounts. The results show that prior to 2015, coefficients fluctuate around zero and are statistically indistinguishable from zero, which provides evidence in support of the parallel trends assumption. In other words, misconduct outcomes in firms with higher ex-ante AI intensity evolve similarly to those in firms with lower intensity in the pre-treatment period.

Following the release of Google TensorFlow in 2015, the coefficients become negative, and the magnitudes increase in subsequent years. This pattern indicates that firms with greater pre-treatment AI exposure experience a relative decline in misconduct outcomes after the exogenous shock, consistent with AI adoption curbing workplace violations. The fact that the effects strengthen over time suggests that the benefits of AI adoption for compliance and monitoring may accumulate gradually as firms adapt and integrate these technologies into their operations.

Taken together, the evidence not only validates the identification strategy by demonstrating pre-treatment parallel trends but also provides richer insights into the dynamics of the treatment effect, reinforcing the interpretation that AI adoption plays a causal role in reducing workplace misconduct.

5.4 Mechanism Tests

We conduct additional analysis to provide corroborating evidence on the mechanism linking AI worker share and workplace misconduct.

First, the effectiveness of investment in the AI workforce in reducing workplace misconduct depends heavily on complementary investments in human capital. For the benefits of AI to materialize, firms need to simultaneously increase discretionary expenditures such as training, monitoring systems, and employee welfare. If this is the operative channel, we should observe that

firms pairing AI adoption with increased discretionary spending on training, monitoring systems, or employee support are better able to embed AI into compliance processes and experience less workplace misconduct.

Second, while we control for employment-related variables, it is possible that the effect of the AI workforce share on workplace misconduct is driven by employee layoffs. If AI automates high-risk or low-skill tasks, firms may observe fewer violations simply because the most accident-prone jobs are replaced by algorithms or robotics. This mechanism implies a mechanical rather than ethical decline in misconduct. Distinguishing this displacement effect from other channels is crucial for interpreting AI's broader ethical implications. If this is the reason we observe less workplace misconduct, and employees prone to accidents are replaced with AI systems operated by AI-proficient workers, we would expect fewer labor violations.

Third, we expect that the effect of AI adoption on workplace misconduct operates primarily through labor productivity. Higher productivity resulting from AI adoption may reduce pressures that typically lead to labor and employment-related violations, such as wage underpayment due to excessive hours or unsafe working conditions. Productivity gains allow firms to meet performance targets without resorting to exploitative practices. Thus, we expect that AI-enhanced productivity will reduce workplace misconduct.

We construct a measure of abnormal discretionary expenses per employee (*AbDiscExpCuts*). A key empirical challenge is that firms are not required to separately disclose safety or employment related expenditures in their financial statements. To address this issue, we follow the procedure of Caskey and Ozel (2017) and Roychowdhury (2006), estimating abnormal discretionary expenses

from the following pooled OLS specification:

$$\frac{SGA_{i,t}}{Emp_{i,t-1}} = \beta_0 + \beta_1 \frac{1}{Emp_{i,t-1}} + \beta_2 \frac{Sales_{i,t-1}}{Emp_{i,t-1}} + \lambda_{j,t} + \alpha_i + \varepsilon_{i,t}, \tag{4}$$

where SGA denotes selling, general, and administrative expenses, Emp is the number of employees, and $Sales$ is net turnover. The specification includes firm fixed effects, industry-year fixed effects, with standard errors two-way clustered at the firm and year level. Following convention, we define *AbDiscExp*$_{i,t}$ as the regression residuals, such that larger values represent greater increase in discretionary expenses.

To examine whether layoffs amplify the effect of AI adoption on workplace misconduct, we construct the variable *Layoff*, defined as the annual growth rate of the total number of employees. For another mechanism check, we measure *LaborProductivity* as sales per employee, expressed in natural logarithm.

Table 6 investigates the mechanisms through which AI adoption affects workplace misconduct by interacting the treatment variable with proxies for abnormal discretionary expenditure increase, layoffs, and labor productivity. The results yield several important insights. We find two main channels though wich AI adoption affects workplace misconduct. First, firms with higher abnormal discretionary expenditures experience a larger decline in workplace violations and penalties after adopting AI. This pattern suggests that greater discretionary spending, often directed towards monitoring, training, and employee-welfare programs, complements AI and amplifies positive compliance with labor law. In other words, when firms invest in these resources, AI's capacity to improve workplace oversight is enhanced, which yields larger reductions in workplace misconduct.

Second, the results highlight the productivity channel. Firms with higher labor productivity

experience sharper post-treatment declines in workplace misconduct. This pattern is consistent with the idea that AI adoption complements existing organizational efficiency: more productive firms may be better positioned to integrate AI technologies into their operations, using them to streamline monitoring processes, detect misconduct more effectively, and reinforce compliance structures. High-productivity firms are also likely to have better managerial and organizational capacity to extract value from AI, thereby amplifying the effect of adoption on misconduct reduction.

We find that layoffs are unlikely to account for the observed decline in workplace misconduct. In Table 6, Panel A and Panel B, the coefficients of the interaction term with layoff-based proxies are statistically insignificant, hence the evidence indicates that workforce reductions do not mediate the post-AI decline in workplace misconduct.

Taken together, these results indicate that AI's effect on workplace misconduct operates mainly through investments in soft skills and productivity improvements, rather than through labor adjustments. The evidence highlights that AI is most effective in reducing misconduct when it is integrated into efficient, well-managed organizations that can leverage its monitoring and compliance enhancing potential.

6 Heterogeneity Analysis and Robustness Checks

In this section, we extend our baseline analysis along two dimensions. First, we explore heterogeneity in the effects of AI adoption on workplace misconduct, and then we assess the robustness of our results through a series of additional checks.

6.1 Heterogeneity Analysis

This section examines whether the effect of AI adoption on workplace misconduct varies systematically across different types of firms. The heterogeneity analysis is important for shedding light on the boundary conditions of our main findings: if the decline in misconduct is concentrated among particular types of firms, this provides insight into when and where AI is most effective as a compliance-enhancing technology.

Panel A of Table 7 shows that the negative association between AI adoption and workplace violations is significantly stronger among larger firms, as well as those with higher levels of intangible capital and organizational capital. Panel B demonstrates a similar pattern for penalty amounts: larger and more intangible-capital-intensive firms experience significantly greater reductions in fines following AI adoption. These findings are corroborated by the dynamic difference-in-differences plots in Figures 2 and 3, which show that the divergence between high- and low-characteristic firms emerges after the 2015 TensorFlow release, while pre-treatment trends remain flat and statistically indistinguishable from zero.

Several underlying reasons may explain these heterogeneous effects. Larger firms tend to face greater regulatory scrutiny, both because they operate across multiple jurisdictions and because their violations often attract higher public and media attention. This heightened visibility raises the reputational and financial costs of misconduct, thereby increasing the incentive to deploy AI tools as part of compliance and monitoring systems. Moreover, larger firms are more frequently targeted by regulators for enforcement actions, which amplifies the benefits of reducing misconduct through enhanced oversight. A further explanation for the stronger effects in large firms lies in the role of economies of scale in AI adoption. Implementing AI systems typically involves substantial

fixed costs: building digital infrastructure, collecting and curating data, and hiring or training specialized staff. Larger firms can spread these fixed costs over a broader asset base, higher revenues, and larger employee populations, which reduces the average cost of adoption. In turn, this scale advantage allows them not only to adopt AI earlier but also to deploy it more comprehensively across multiple divisions and geographies.

The stronger effects among firms with greater intangible and organizational capital are consistent with the idea that AI is a technology that complements, rather than substitutes for, pre-existing organizational capabilities. Firms rich in intangible assets (e.g. R&D, patents, or proprietary processes) are likely to have more advanced digital infrastructure and human capital that facilitate the integration of AI into compliance and oversight functions. Similarly, firms with stronger organizational capital (e.g., well-established managerial structures, routines, and internal controls) may be better equipped to deploy AI in ways that reinforce governance and reduce misconduct. In contrast, firms lacking such foundations may struggle to adapt AI tools effectively, limiting the impact on misconduct outcomes.

The heterogeneity analysis highlights that AI adoption is not uniformly effective across firms. Instead, its compliance-enhancing benefits are concentrated among firms with greater scale and stronger intangible or organizational assets. This finding matters because it suggests that AI may widen disparities across firms: large, intangible-intensive organizations can leverage AI to improve compliance and reduce misconduct, while smaller or less developed firms may not experience comparable benefits. Such heterogeneity has implications for both regulators and policymakers, as it suggests that the diffusion of AI could unevenly alter the landscape of corporate accountability.

6.2 Other Offense Types

Our main analysis focuses on safety and employment related workplace misconduct, as these categories are most plausibly linked to the channels through which AI adoption operates. Safety violations, such as breaches of occupational health standards, and employment related offenses, such as wage and hour violations, are directly connected to the monitoring and oversight of workplace conditions. These are precisely the domains where AI technologies through enhanced data collection, anomaly detection, and real-time monitoring can plausibly reduce misconduct by improving compliance and lowering information frictions. By contrast, offenses in other categories (e.g., financial reporting violations, competition-related misconduct, or government contracting irregularities) are less likely to be directly mitigated by AI monitoring at the workplace level, and may be influenced by different governance structures and regulatory regimes.

To ensure that our results are not spuriously driven by unobserved factors, we conduct robustness checks by re-estimating our difference-in-differences models for a broad set of other offense types, including competition, consumer protection, environmental, financial, government contracting, and healthcare-related violations. The results, reported in Table 8, show that the interaction between post-2015 AI adoption and pre-treatment AI intensity yields no systematic or robust effects for these alternative categories. Similarly, the dynamic estimates presented in Figure 4 confirm the absence of significant pre- or post-treatment differences in these offense types.

The lack of effects for other offense categories strengthens the interpretation of our main findings in two ways. First, it validates our research design by demonstrating that the observed treatment effects are not a general artifact of AI adoption reducing all forms of corporate misconduct. Instead, the effects are concentrated in the categories most closely related to workplace conditions,

consistent with the hypothesized compliance enhancing channel of AI. Second, it rules out the concern that our results might reflect broader changes in firms' overall governance or reporting behavior. If such changes were driving the results, we would expect to observe reductions across a wider set of offense categories.

These robustness checks provide strong support for our empirical approach and the underlying hypothesis. By showing that AI adoption specifically reduces safety- and employment-related misconduct, while leaving unrelated offense categories unaffected, we demonstrate that the documented effects are both causal and mechanism-consistent.

6.3 Placebo Test

This section evaluates the validity of our identification strategy by designing a placebo test. The core idea behind this exercise is to assess whether the observed post-2015 declines in workplace misconduct are genuinely attributable to the release of Google TensorFlow, or whether similar effects could be obtained if we artificially assign the treatment to an earlier, arbitrary year. If our empirical framework were merely capturing spurious correlations or broader time trends, we would expect to observe "effects" around the placebo treatment year as well.

The placebo design follows the same specification as our main dynamic difference-in-differences, but instead of using 2015 as the treatment year, we assign 2013 as a false treatment year, with 2012 serving as the baseline. This methodology replicates the structure of our main analysis but deliberately misaligns the timing of the treatment, allowing us to test for spurious effects.

The results, shown in Figure 5, reveal no evidence of systematic treatment effects around the placebo year. In both panels (a) for violations and (b) for penalty amounts the estimated coeffi-

cients fluctuate around zero and remain statistically insignificant throughout almost all the placebo pre- and post-treatment period. Equally important, the pre-2013 coefficients also remain flat, indicating no signs of differential pre-trends between high and low AI-intensity firms. Overall, the placebo test demonstrates that our findings are robust and that the observed reductions in workplace misconduct are causally linked to the 2015 release of TensorFlow rather than spurious correlations.

6.4 Alternative Empirical Specifications

We next assess the robustness of our findings using alternative empirical specifications that vary the fixed-effects structure and incorporate linear time trends.

In Table 9, we check the stability of our estimates to alternative fixed-effects structures. Columns (1)–(5) progressively add controls and absorb higher-dimensional heterogeneity: starting from a parsimonious specification without fixed effects or controls (Column (1)), then adding the control variables (Column (2)), only year fixed effects (Column (3)), only firm fixed effects (Column (4)), and finally both firm and year fixed effects (Column (5)).[8] Across all specifications and both outcomes, the treatment effect remains large, negative, and statistically significant, with magnitudes that become more pronounced as we tighten identification with firm and year fixed effects. This pattern is consistent with the baseline estimates.

In addition to our baseline dynamic difference-in-differences design, we also estimate a series of alternative specifications incorporating linear time trends at the industry and state levels. The rationale for these additional tests is twofold. First, while our main difference-in-differences framework already includes firm, and industry-by-year fixed effects, it is possible that certain in-

[8]These specifications also include the firms excluded from the benchmark analysis due to missing state information.

dustries or states exhibit systematic differential trends in workplace misconduct unrelated to AI adoption. For instance, regulatory enforcement intensity may evolve differently across states, or industry-specific shifts in safety standards and compliance practices may generate gradual changes in misconduct outcomes. If such differential trends are not adequately controlled for, our baseline estimates could be biased. Incorporating linear industry- and state-specific time trends therefore provides a more stringent robustness test by absorbing gradual, location- or sector-specific changes that might otherwise confound the treatment effect.

Figure 6 presents the results from these alternative specifications, contrasting models with and without linear trends. Across all panels, the results remain consistent with our baseline findings: firms with greater pre-treatment AI intensity experience sharp and persistent declines in both violations and penalty amounts after 2015. The introduction of industry- and state-specific trends does not materially alter the direction, timing, or magnitude of the treatment effects.

6.5 Alternative AI Measure

As an additional robustness exercise, we re-estimate our baseline models using the narrow definition of AI skills proposed by Babina et al. (2024). Unlike the broad definition employed in the main analysis, the narrow measure excludes skills that are only tangentially related to AI and thus captures a more restrictive set of AI expertise.

The results, presented in Table 10, indicate that the negative relationship between AI adoption and workplace misconduct remains qualitatively similar when using this narrower proxy. Specifically, firms with higher pre-treatment exposure to narrowly defined AI skills continue to experience post-2015 declines in both violations (statistically significant in the fully specified model),

and penalties, although the magnitudes of the some estimates are attenuated relative to those obtained with the broad measure. This finding suggests that our results are not an artifact of the particular AI skill definition employed. Whether AI adoption is measured broadly or narrowly, we find consistent evidence that greater reliance on AI talent reduces workplace misconduct.

6.6 Additional Controls: Corporate Governance and Executive Incentives

A potential concern is that our baseline difference-in-differences estimates may partially capture contemporaneous changes in firms' internal governance and managerial incentives that co-move with AI adoption. Board structure, monitoring intensity, and pay sensitivity to performance and risk (e.g., *Delta* and *Vega*) plausibly shape both compliance-related outcomes and penalty exposure. To ensure that the AI effect is not an artifact of these correlated shifts, we augment the baseline specification with a rich set of governance and compensation controls.

Table 11 reports the results. The post$\times$AI exposure coefficient remains large, negative, and statistically significant across five of six specifications for violations and penalties. Magnitudes are close to the benchmark without these controls, indicating that governance and compensation do not explain away the treatment effect. In sum, controlling for an extensive set of governance and incentive variables leaves the main DiD effect intact in sign, and magnitude. The stability of coefficients strengthens the interpretation that the measured decline in violations and penalties is tied to pre-2015 AI exposure rather than contemporaneous shifts in board composition or executives' risk-incentive profiles.

Figure 7 presents the dynamic difference-in-differences estimates corresponding to the specification that includes the extended set of control variables. The results closely mirror those from

the benchmark specification, with the dynamic treatment effects remaining consistently negative and statistically significant in the post-2014 period. The absence of significant pre-treatment coefficients further supports the validity of the parallel-trends assumption. Overall, these patterns indicate that our main findings are not driven by omitted firm-level characteristics related to governance or executive incentives, reaffirming the robustness of the causal link between AI adoption and reductions in workplace misconduct.

6.7 Including Excluded Industries: Finance, Utilities, and Other

Our benchmark sample excludes firms in the Finance, Utilities, and Other industries since they operate under distinct regulatory frameworks and exhibit fundamentally different patterns of reporting, compliance, and risk exposure. In this section, we assess the robustness of our findings to including these industries. Table 12 reports the corresponding DiD estimates, and Figure 8 presents the associated dynamic difference-in-differences specification.

The inclusion of these sectors does not materially alter our main conclusions. The estimated coefficients on the interaction term between the post-2014 period and firm-level AI adoption remain negative and statistically significant across all specifications, both for the number and monetary value of violations. The dynamic difference-in-differences estimates in Figure 8 confirm that the timing and magnitude of the treatment effects are virtually unchanged, with no evidence of differential pre-trends.

This robustness exercise suggests that our results are not driven by sample selection related to the exclusion of highly regulated or structurally distinct industries. Rather, the observed reduction in workplace misconduct following AI adoption reflects a broad and systematic pattern across the

corporate sector.

6.8 Propensity Score Matching

To further address concerns related to selection on observables and potential differences in pre-treatment characteristics between AI-intensive and non–AI-intensive firms, we implement a propensity score matching strategy as a robustness exercise. The objective of this approach is to construct a comparison group of firms that are observationally similar to treated firms prior to the AI shock, thereby strengthening the credibility of the parallel trends assumption underlying our DiD design.

We estimate propensity scores using a logit specification in the pre-treatment year (2014), where treatment status is defined based on firms' exposure to AI prior to the release of Google TensorFlow. The matching model includes a rich set of lagged firm characteristics commonly used in the literature to capture financial structure, scale, investment behavior, and innovation capacity such as leverage, cash flow, dividend payouts, firm size, sales, employment, labor intensity, asset turnover, market-to-book ratio, asset tangibility, capital expenditures, firm age, and R&D intensity.

Panel estimates based on the propensity-score-matched sample are reported in Table 13. The estimated coefficients on the interaction between *Post* and the pre-treatment share of AI workers remain negative, statistically significant, and economically meaningful across all outcome measures. The magnitude of the effects is comparable to those obtained in the baseline specifications. This finding suggests that the main results are not driven by systematic differences in observable firm characteristics.

Figure 9 further corroborates these findings by presenting dynamic DiD estimates based on the matched sample. The coefficients show no evidence of differential pre-trends prior to 2015,

while a sharp and persistent decline in both violations and penalty amounts emerges following the AI shock. Taken together, the propensity score matching results reinforce our interpretation that the observed reduction in workplace misconduct reflects a causal effect of AI adoption rather than compositional changes or selection into treatment.

6.9 Controlling for Other Firm-level Technologies

A potential concern is that the estimated effects attributed to AI adoption may in fact capture broader patterns of technological upgrading or digitalization at the firm level. To address this concern, we conduct an additional robustness exercise that explicitly controls for other firm-level technologies, following the classification proposed by Babina et al. (2024). This approach allows us to disentangle the effect of AI adoption from contemporaneous investments in related but conceptually distinct-technologies.

Specifically, we augment our baseline DiD specification by including firm-level measures of exposure to robotics, information technology (IT), data infrastructure, and data analytics. These variables capture alternative technological channels that may influence firm behavior, monitoring capacity, or operational efficiency, and thus could independently affect workplace misconduct. Importantly, while these technologies are often correlated with AI adoption, they differ in their economic role such that robotics primarily affects physical automation, IT reflects general digital infrastructure, and data-related technologies capture data storage and processing capabilities rather than algorithmic decision-making.

Table 14 reports the results from this augmented specification. The estimated coefficient on the interaction between *Post* and the pre-treatment share of AI workers remains negative, highly

statistically significant, and economically large across all outcome measures. The magnitude of the AI effect is largely unchanged relative to the baseline estimates, indicating that our main results are not driven by firms' adoption of other technologies. By contrast, the coefficients on robot share, IT share, data share, and data analytics share are generally small and statistically insignificant, suggesting that these technologies do not independently account for the observed reduction in violations or penalty amounts.

Figure 10 further reinforces this interpretation by presenting dynamic DiD estimates that condition on the full set of technology controls. The coefficients exhibit no evidence of differential pre-trends prior to the 2015 TensorFlow release, while a sharp and persistent post-treatment decline emerges only for AI-exposed firms. This pattern supports the view that AI adoption rather than generic technological upgrading drives the observed changes in workplace misconduct.

Taken together, these findings highlight the distinct role of AI as a general-purpose technology with unique monitoring, prediction, and decision-support capabilities. While firms may simultaneously invest in multiple technologies, the results indicate that AI has an independent and robust effect that cannot be explained by broader digitalization or automation trends.

7 Conclusion

This paper examines the relationship between artificial intelligence (AI) adoption, and workplace misconduct. While the broader literature has emphasized AI's implications for productivity, innovation, and firm value, we focus on its consequences for workplace misconduct that has important economic and social implications.

To identify the causal effect of AI adoption on workplace misconduct, we exploit the 2015

release of Google TensorFlow as a plausibly exogenous shock within a difference-in-differences framework. Our analysis delivers four key findings. First, we show that firms with greater AI intensity are associated with fewer safety and employment related violations and lower penalties, and these patterns remain robust across multiple regression specifications. Second, our causal evidence shows that firms with higher pre-treatment AI intensity experience sharper and persistent declines in workplace misconduct in the years following the shock. Third, our mechanism analysis reveals that these reductions are driven by productivity-enhancing complementarities and increases in discretionary expenditures that potentially improve the AI literacy of the workforce. Importantly, we find no evidence that misconduct declines arise from labor-adjustment channels such as layoffs. Fourth, heterogeneity analysis shows that the compliance benefits of AI are concentrated among large, intangible- and organizational-capital intensive firms, reflecting economies of scale and stronger capacities to integrate AI into governance and oversight structures.

These results have two broader implications. At the firm level, they suggest that AI adoption can improve organizational compliance and reduce exposure to regulatory risk, offering benefits that extend beyond productivity gains. At the societal level, they highlight that AI adoption generates positive externalities by curbing workplace misconduct, particularly in safety and employment related domains. Yet, the uneven distribution of benefits highlights a potential widening gap between larger, intangible-rich firms and their smaller counterparts. Policymakers may therefore wish to encourage broader diffusion of AI-enabled compliance technologies to ensure more equitable improvements in workplace standards as otherwise AI may widen compliance inequality. This can be done through AI such as open-source compliance tools, public–private partnerships, or targeted subsidies that would help to equalize the benefits of AI adoption.

In sum, our findings position AI adoption as a novel determinant of workplace conduct, op-

erating primarily through productivity and investment channels rather than labor adjustment. By documenting both its compliance benefits and heterogeneity, this paper broadens our understanding of the economic and social consequences of AI adoption and provides a foundation for future research on the intersection of technology, governance, and employee welfare.

References

Akyildirim, E., G. O. Danisman, and S. Ongena (2025). AI employment and political risk disclosures in earnings calls. *Swiss Finance Institute Research Paper* (25-56).

Ashraf, M. (2025). Does automation improve financial reporting? Evidence from internal controls. *Review of Accounting Studies 30*(1), 436–479.

Autor, D. H. (2015). Why are there still so many jobs? The history and future of workplace automation. *Journal of Economic Perspectives 29*(3), 3–30.

Babina, T., A. Fedyk, A. He, and J. Hodson (2024). Artificial intelligence, firm growth, and product innovation. *Journal of Financial Economics 151*, 103745.

Becker, G. S. (1968). Crime and punishment: An economic approach. *Journal of Political Economy 76*(2), 169–217.

Brynjolfsson, E., B. Chandar, and R. Chen (2025). Canaries in the coal mine? Six facts about the recent employment effects of artificial intelligence. *Stanford Digital Economy Lab. Published August*.

Brynjolfsson, E., A. Korinek, and A. K. Agrawal (2025). A research agenda for the economics of transformative AI.

Brynjolfsson, E., D. Li, and L. Raymond (2025). Generative AI at work. *The Quarterly Journal of Economics 140*(2), 889–942.

Caskey, J. and N. B. Ozel (2017). Earnings expectations and employee safety. *Journal of Accounting and Economics 63*(1), 121–141.

Chen, W. and S. Srinivasan (2024). Going digital: Implications for firm value and performance. *Review of Accounting Studies 29*(2), 1619–1665.

Chen, Y., E. Ofosu, D. O'Sullivan, M. Veeraraghavan, and L. Zolotoy (2025). Rank-and-file employee stock options and workplace safety. *Management Science 71*(7), 5971–5996.

Chircop, J., M. Tarsalewska, and A. Trzeciakiewicz (2025). CEO risk taking equity incentives and workplace misconduct. *The Accounting Review 100*(1), 139–167.

Christensen, D. M. (2016). Corporate accountability reporting and high-profile misconduct. *The Accounting Review 91*(2), 377–399.

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.

Ciconte, W., A. Rozario, and O. Urcan (2024a). Artificial intelligence regulation and investor risk: Evidence from state and local artificial intelligence mandates. *Available at SSRN 5023685*.

Ciconte, W., A. Rozario, and O. Urcan (2024b). Do AI laws inhibit innovation? *Available at SSRN 5046045*.

Cockburn, I. M., R. Henderson, and S. Stern (2018). The impact of artificial intelligence on innovation: An exploratory analysis. In *The economics of artificial intelligence: An agenda*, pp. 115–146. University of Chicago Press.

Cohn, J. B. and M. I. Wardlaw (2016). Financing constraints and workplace safety. *The Journal of Finance 71*(5), 2017–2058.

Coles, J. L., N. D. Daniel, and L. Naveen (2006). Managerial incentives and risk-taking. *Journal of Financial Economics 79*(2), 431–468.

Coles, J. L., N. D. Daniel, and L. Naveen (2013). Calculation of compensation incentives and firm-related wealth using Execucomp: Data, program, and explanation. *Available at SSRN 2296381*.

Demirci, O., J. Hannane, and X. Zhu (2025). Who is AI replacing? The impact of generative AI on online freelancing platforms. *Management Science*.

Ding, K., B. Lev, X. Peng, T. Sun, and M. A. Vasarhelyi (2020). Machine learning improves accounting estimates: Evidence from insurance payments. *Review of Accounting Studies 25*(3), 1098–1134.

Eloundou, T., S. Manning, P. Mishkin, and D. Rock (2024). GPTs are GPTs: Labor market impact potential of LLMs. *Science 384*(6702), 1306–1308.

Fedyk, A., J. Hodson, N. Khimich, and T. Fedyk (2022). Is artificial intelligence improving the audit process? *Review of Accounting Studies 27*(3), 938–985.

Felten, E., M. Raj, and R. Seamans (2021). Occupational, industry, and geographic exposure to artificial intelligence: A novel dataset and its potential uses. *Strategic Management Journal 42*(12), 2195–2217.

Gihleb, R., O. Giuntella, L. Stella, and T. Wang (2022). Industrial robots, workers' safety, and health. *Labour Economics 78*, 102205.

Gmyrek, P., J. Berg, and D. Bescond (2023). Generative AI and jobs: A global analysis of potential effects on job quantity and quality. *ILO Working paper 96*.

Grennan, J. and K. Li (2023). Corporate culture: A review and directions for future research. *Handbook of Financial Decision Making*, 112–132.

Handa, K., A. Tamkin, M. McCain, S. Huang, E. Durmus, S. Heck, J. Mueller, J. Hong, S. Ritchie, T. Belonax, et al. (2025). Which economic tasks are performed with AI? Evidence from millions of Claude conversations. *arXiv preprint arXiv:2503.04761*.

Heese, J. and J. Pacelli (2025). Enterprise resource planning (ERP) system implementations and corporate misconduct. *The Accounting Review 100*(1), 291–315.

Heese, J. and G. Pérez-Cavazos (2020). When the boss comes to town: The effects of headquarters' visits on facility-level misconduct. *The Accounting Review 95*(6), 235–261.

Hui, X., O. Reshef, and L. Zhou (2024). The short-term effects of generative artificial intelligence on employment: Evidence from an online labor market. *Organization Science 35*(6), 1977–1989.

International Labour Organization (2025). Occupational safety and health - A guide for labour inspectors and other stakeholders: Introduction and acknowledgements. Retrieved from https://www.ilo.org/topics-and-sectors/occupational-safety-and-health-guide-labour-inspectors-and-other/introduction-and-acknowledgements.

Li, X. and A. Raghunandan (2021). Institutional ownership and workplace misconduct: Evidence from federal labor law violations. *Available at SSRN 3460126*.

Liang, C. Y., Y. Qi, R. A. Zhang, and H. Zhu (2023). Does sunlight kill germs? Stock market listing and workplace safety. *Journal of Financial and Quantitative Analysis 58*(4), 1645–1674.

Peters, R. H. and L. A. Taylor (2017). Intangible capital and the investment-q relation. *Journal of Financial Economics 123*(2), 251–272.

Raghunandan, A. (2021). Financial misconduct and employee mistreatment: Evidence from wage theft. *Review of Accounting Studies 26*(3), 867–905.

Raghunandan, A. and S. Rajgopal (2024). Do socially responsible firms walk the talk? *The Journal of Law and Economics 67*(4), 767–810.

Raghunandan, A. and T. G. Ruchti (2024). The impact of information frictions within regulators: Evidence from workplace safety violations. *Journal of Accounting Research 62*(3), 1067–1120.

Rock, D. (2019). Engineering value: The returns to technological talent and investments in artificial intelligence. *Available at SSRN 3427412*.

Roychowdhury, S. (2006). Earnings management through real activities manipulation. *Journal of Accounting and Economics 42*(3), 335–370.

Tomlinson, K., S. Jaffe, W. Wang, S. Counts, and S. Suri (2025). Working with AI: Measuring the occupational implications of generative AI. *arXiv preprint arXiv:2507.07935*.

United States Department of Labor, Wage and Hour Division (2024). Artificial intelligence and automated systems in the workplace under the fair labor standards act and other federal labor standards. https://data.aclum.org/storage/2025/01/DOL_www_dol_gov_sites_dolgov_files_WHD_fab_fab2024_1.pdf. Field Assistance Bulletin No. 2024-1.

United States Senate Committee on Health, Education, Labor, and Pensions (HELP) (2024). The

"injury-productivity trade-off": How amazon's obsession with speed creates uniquely dangerous warehouses. https://www.help.senate.gov/imo/media/doc/amazon_investigation.pdf. Majority Staff Report, December 2024.

Table 1: Sample Composition by Year and Industry

Panel A: Sample Composition by Year		
Year	Number of Observations	Percentage of Observations
2010	715	9.0
2011	753	9.5
2012	780	9.9
2013	821	10.4
2014	871	11.0
2015	842	10.7
2016	824	10.4
2017	799	10.1
2018	781	9.9
2019	720	9.1
Total	7,906	100.0
Panel B: Sample Composition by Industry		
Industry	Number of Observations	Percentage of Observations
Consumer Non-Durables	246	3.1
Consumer Durables	341	4.3
Manufacturing	1,130	14.3
Oil, Gas, and Coal	66	0.8
Chemicals	365	4.6
High Tech	2,557	32.3
Telephone and Television	128	1.6
Wholesale and Retail	1,002	12.7
Healthcare	1,694	21.4
Other	377	4.8
Total	7,906	100.0

Note: This table summarizes the sample distribution by year (Panel A) and by Fama–French industry classification (Panel B).

Table 2: Summary Statistics

	Mean	SD	P25	P50	P75	Min	Max	Count
Violations	0.279	0.794	0.000	0.000	0.000	0.000	6.000	7,906
Log Violations	0.149	0.375	0.000	0.000	0.000	0.000	1.946	7,906
Penalty	6,386.78	53,396.32	0.000	0.000	0.000	0.000	1,819,730.00	7,806
Log Penalty	1.368	3.366	0.000	0.000	0.000	0.000	14.414	7,806
Leverage	0.534	0.374	0.316	0.496	0.668	0.031	8.147	7,888
CashFlow	-0.008	0.362	-0.004	0.072	0.116	-9.674	0.433	7,880
DividendPayout	0.010	0.021	0.000	0.000	0.014	0.000	0.258	7,880
Size	6.364	1.917	5.172	6.468	7.699	-0.921	11.080	7,886
Sales	0.108	0.476	-0.031	0.048	0.148	-1.000	9.550	7,850
Employees	9.966	21.180	0.557	2.500	9.000	0.000	204.700	7,881
LaborIntensity	8.690	12.045	2.799	4.986	9.402	0.000	180.375	7,881
AssetTurnover	1.043	0.692	0.598	0.888	1.331	0.000	5.910	7,899
Market-to-Book	2.119	1.991	1.031	1.565	2.553	0.224	77.226	7,727
Tangibility	0.184	0.168	0.062	0.131	0.252	0.000	0.918	7,902
Capex	0.041	0.043	0.016	0.029	0.051	0.000	0.609	7,899
R&D Intensity	0.093	0.154	0.008	0.037	0.117	0.000	3.269	7,855
FirmAge	3.404	0.621	2.996	3.434	3.871	1.099	4.779	7,901
Number of Directors	2.240	0.213	2.079	2.303	2.398	1.609	2.708	5,502
Non-Executive Director Ratio	0.091	0.024	0.074	0.086	0.102	0.030	0.188	5,505
Log Delta	3.745	1.202	2.968	3.788	4.575	-0.304	7.121	4,971
Log Vega	2.438	1.579	1.221	2.683	3.646	0.000	6.149	5,013
Log FirmRelatedWealth	8.091	1.154	7.335	8.133	8.877	4.061	11.378	4,965
Log CEOtenure	1.854	0.805	1.320	1.946	2.485	0.000	3.611	4,809
Log CEOage	3.972	0.077	3.930	3.970	4.025	3.714	4.190	4,971
Log Labor Productivity	-3.762	1.980	-5.065	-3.660	-2.411	-9.763	2.675	5,132
Log Total Intangible Capital	5.784	1.600	4.700	5.708	6.921	0.523	10.335	7,886
Log Organizational Capital	4.464	1.871	3.436	4.562	5.732	0.000	9.016	7,905
Share of AI Workers	0.006	0.005	0.002	0.004	0.007	0.000	0.044	6,741

Note: This table reports descriptive statistics for variables used in the analysis. All nominal variables are deflated using the Consumer Price Index (CPI) 1990 U.S. dollars.

Table 3: Reduced Form Regression Estimates

	(1) Log Violations	(2) Asinh Violations	(3) PPML Violations	(4) Log Penalty	(5) Asinh Penalty	(6) PPML Q(Penalty)
L.Share of AI Workers	-1.653*	-2.153*	-56.14**	-8.825	-9.511	-4.325
	(0.868)	(1.125)	(26.56)	(8.342)	(8.939)	(5.587)
Leverage	0.00829	0.0110	0.539*	0.104	0.110	0.0690
	(0.0108)	(0.0140)	(0.318)	(0.111)	(0.119)	(0.0727)
CashFlow	0.000269	0.000616	-0.0196	0.0218	0.0228	0.0358
	(0.00995)	(0.0129)	(0.667)	(0.0990)	(0.106)	(0.0636)
DividendPayout	0.326	0.427	1.604	3.424	3.656	1.310
	(0.231)	(0.297)	(1.815)	(2.609)	(2.809)	(0.993)
Size	0.0118	0.0152	0.0990	0.139**	0.148**	0.0511
	(0.00719)	(0.00935)	(0.105)	(0.0627)	(0.0671)	(0.0321)
Sales	-0.000749	-0.00106	-0.00608	0.00419	0.00477	0.0139
	(0.00242)	(0.00314)	(0.123)	(0.0234)	(0.0251)	(0.0157)
Employment	-0.000137	-0.000177	-0.00103	-0.0102	-0.0112	-0.00388*
	(0.00118)	(0.00154)	(0.00396)	(0.0110)	(0.0117)	(0.00227)
LaborIntensity	-0.000591	-0.000788	-0.0123	0.00118	0.00110	-0.000642
	(0.00131)	(0.00170)	(0.0134)	(0.0129)	(0.0138)	(0.00582)
AssetTurnover	-0.00692	-0.00887	-0.0417	-0.0416	-0.0458	-0.0269
	(0.0108)	(0.0141)	(0.122)	(0.103)	(0.111)	(0.0486)
Market-to-Book	-0.000448	-0.000589	0.0444	-0.00182	-0.00171	0.00484
	(0.00201)	(0.00260)	(0.0713)	(0.0197)	(0.0212)	(0.0131)
Tangibility	0.0323	0.0437	0.374	0.376	0.396	0.132
	(0.0868)	(0.113)	(0.714)	(0.750)	(0.802)	(0.298)
Capex	-0.00634	-0.00457	0.650	-0.123	-0.137	-0.0921
	(0.0952)	(0.123)	(1.234)	(0.845)	(0.906)	(0.428)
FirmAge	-0.00742	-0.00833	0.299	-0.427	-0.461	-0.0609
	(0.0515)	(0.0672)	(0.272)	(0.418)	(0.446)	(0.130)
R&D Intensity	0.00188	0.00222	-4.051	0.0427	0.0436	-0.0101
	(0.0159)	(0.0207)	(3.225)	(0.137)	(0.147)	(0.0847)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Industry-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.527	0.527	–	0.473	0.473	–
Observations	7,229	7,229	2,441	7,170	7,170	7,170

Note: This table reports panel regression estimates of the relationship between firm-level workplace misconduct outcomes and the lagged share of AI workers. Columns (1) and (4) use the $ln(x+1)$ transformation of the dependent variable, columns (2) and (5) use the inverse hyperbolic sine ($asinh(x)$) transformation, and columns (3) and (6) employ Poisson pseudo–maximum likelihood with high-dimensional fixed effects. For column (6), we construct a categorical variable based on the deciles of the total penalty value within each industry (2-digit *NAICS*) and year. The number of observations in the PPML estimation for the number of violations is smaller because the PPML estimator drops separated observations, cases in which the fixed effects create disconnected groups of observations. Since such separated observations lead to convergence issues, the estimator automatically excludes them from the sample. The similar principle applies to the other PPML estimation results for the number of violations. Standard errors are clustered at the firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: DiD Estimates: Identification Check Using Share of AI Workers

	(1)	(2)	(3)	(4)	(5)	(6)
	Share of AI Workers	Share of AI Workers	Share of AI Workers	Share of AI Workers	Share of AI Workers	Share of AI Workers
Post × Share of AI Workers$_{2014}$	0.255**	0.416***	0.374**	0.369**	0.308**	0.288**
	(0.0901)	(0.113)	(0.122)	(0.121)	(0.107)	(0.108)
Control Variables	No	Yes	Yes	Yes	Yes	Yes
Industry-Year FE	No	No	Yes	Yes	Yes	Yes
Firm FE	No	No	No	No	Yes	Yes
Time Trends	No	No	No	No	No	Yes
Adjusted R^2	0.646	0.682	0.692	0.697	0.775	0.775
Observations	14,300	3,058	3,029	3,020	2,993	2,993

Note: This table reports Difference-in-Differences (DiD) estimates of the effect of firm-level AI adoption on the firm-level share of AI workers. Control variables include a lag of Leverage, CashFlow, DividendPayout, FirmSize, Sales, Employees, LaborIntensity, AssetTurnover, Market-to-Book, Tangibility, Capex, R&D Intensity, and FirmAge; as well as governance variables including Vega, Delta, FirmRelatedWealth, CEO age, CEO tenure, board size, and the ratio of non-executive directors. The trends are state-specific and industry-specific linear time trends. Standard errors (in parentheses) are two-way clustered at the firm and year level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: DiD Estimates: Baseline Results

	(1) Log Violations	(2) Asinh Violations	(3) PPML Violations	(4) Log Penalty	(5) Asinh Penalty	(6) PPML Q(Penalty)
Post × Share of AI Workers$_{2014}$	**-6.992*****	**-9.077*****	**-122.3*****	**-58.88*****	**-63.09*****	**-27.66*****
	(1.443)	(1.870)	(26.89)	(14.75)	(15.74)	(6.459)
Leverage	0.00935	0.0123	0.523***	0.0954	0.102	0.0771
	(0.00799)	(0.0103)	(0.174)	(0.108)	(0.116)	(0.0709)
CashFlow	-0.00216	-0.00257	-0.392	0.0228	0.0240	0.0378
	(0.00890)	(0.0115)	(0.613)	(0.0932)	(0.100)	(0.0589)
DividendPayout	0.186	0.239	1.396	2.640	2.847	1.285
	(0.228)	(0.294)	(2.081)	(2.669)	(2.856)	(0.879)
Size	0.0130**	0.0168**	0.175	0.130**	0.139**	0.0613**
	(0.00527)	(0.00682)	(0.109)	(0.0447)	(0.0481)	(0.0274)
Sales	-0.000792	-0.00106	-0.00822	-0.00572	-0.00588	0.00749
	(0.00178)	(0.00231)	(0.135)	(0.0203)	(0.0217)	(0.0155)
Employment	0.000698	0.000910	-0.0000627	0.0000697	-0.000117	-0.00193
	(0.000907)	(0.00120)	(0.00334)	(0.00522)	(0.00544)	(0.00118)
LaborIntensity	-0.000160	-0.000223	-0.00106	0.00157	0.00161	0.00207
	(0.00116)	(0.00151)	(0.0153)	(0.0135)	(0.0144)	(0.00596)
AssetTurnover	-0.00594	-0.00761	-0.0305	-0.0537	-0.0582	-0.0322
	(0.00905)	(0.0118)	(0.105)	(0.0621)	(0.0666)	(0.0341)
Market-to-Book	-0.000286	-0.000379	0.0330	0.00594	0.00650	0.00518
	(0.00161)	(0.00209)	(0.0529)	(0.0137)	(0.0147)	(0.00877)
Tangibility	-0.0309	-0.0389	0.205	-0.537	-0.575	-0.131
	(0.0711)	(0.0941)	(0.716)	(0.549)	(0.583)	(0.176)
Capex	-0.00987	-0.00935	1.289	-0.367	-0.396	-0.205
	(0.0612)	(0.0796)	(1.333)	(0.436)	(0.463)	(0.226)
FirmAge	-0.0232	-0.0296	0.191	-0.521	-0.561	-0.0795
	(0.0332)	(0.0434)	(0.238)	(0.320)	(0.342)	(0.117)
R&D Intensity	0.00758	0.00970	-7.024***	0.116	0.122	0.0114
	(0.0130)	(0.0170)	(2.170)	(0.117)	(0.125)	(0.0669)
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes
Industry-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.529	0.529	–	0.482	0.482	–
Observations	7,906	7,906	2,798	7,842	7,842	7,895

Note: This table reports Difference-in-Differences (DiD) estimates of the effect of firm-level AI adoption on workplace misconduct. *Post* equals one for $t \geq 2015$ (Google TensorFlow release). The 2014 share of AI workers proxies for pre-treatment AI intensity. Columns (1) and (4) use the $ln(x+1)$ transformation, columns (2) and (5) use the inverse hyperbolic sine ($asinh(x)$) transformation, and columns (3) and (6) employ Poisson pseudo–maximum likelihood with high-dimensional fixed effects. For column (6), we construct a categorical variable based on the deciles of the total penalty value within each industry (2-digit *NAICS*) and year. Standard errors (in parentheses) are two-way clustered at the firm and year level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: DiD Estimates: Mechanism Tests

Panel A: Dependent Variable – Log Violations			
	(1)	(2)	(3)
Post × Share of AI Workers$_{2014}$	-7.444***	-7.495***	9.978**
	(1.633)	(1.561)	(4.405)
Post × Share of AI Workers$_{2014}$ × L.AbDiscExp	-43.34*		
	(19.41)		
Post × Share of AI Workers$_{2014}$ × L.Layoff		-0.149	
		(3.393)	
Post × Share of AI Workers$_{2014}$ × L.LaborProductivity			-3.991***
			(1.054)
Control Variables	Yes	Yes	Yes
Industry-Year FE	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes
Adjusted R^2	0.520	0.525	0.526
Observations	6,513	7,510	7,862
Panel B: Dependent Variable – Log Penalty Amount			
	(1)	(2)	(3)
Post × Share of AI Workers$_{2014}$	-62.69***	-61.02***	58.79
	(16.32)	(16.38)	(38.43)
Post × Share of AI Workers$_{2014}$ × L.AbDiscExp	-414.8**		
	(136.4)		
Post × Share of AI Workers$_{2014}$ × L.Layoff		-22.93	
		(35.79)	
Post × Share of AI Workers$_{2014}$ × L.LaborProductivity			-27.64**
			(9.161)
Control Variables	Yes	Yes	Yes
Industry-Year FE	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes
Adjusted R^2	0.468	0.478	0.480
Observations	6,456	7,448	7,802

Note: This table reports Difference-in-Differences (DiD) estimates examining the mechanisms through which firm-level AI adoption affects workplace misconduct. Panel A focuses on the natural logarithm of one plus the number of violations, while Panel B examines the natural logarithm of one plus the penalty amounts. Interaction terms capture mechanism effects through abnormal discretionary expenses (L.AbDiscExp), layoffs (L.Layoff), and labor productivity (L.LaborProductivity). Control variables include a *lag* of Leverage, CashFlow, DividendPayout, FirmSize, Sales, Employees, LaborIntensity, AssetTurnover, Market-to-Book, Tangibility, Capex, R&D Intensity, and FirmAge. The specification includes industry-by-year, and firm fixed effects. Standard errors are two-way clustered at the firm and year level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 7: DiD Estimates: Heterogeneity Analysis

Panel A: Dependent Variable – Log Violations			
	(1)	(2)	(3)
Post × Share of AI Workers$_{2014}$	12.63**	8.765	4.131
	(4.430)	(5.323)	(3.573)
Post × Share of AI Workers$_{2014}$ × L.Log Assets	-3.440***		
	(0.824)		
Post × Share of AI Workers$_{2014}$ × L.Intangible Capital		-2.975**	
		(0.955)	
Post × Share of AI Workers$_{2014}$ × L.Organizational Capital			-2.734***
			(0.770)
Control Variables	Yes	Yes	Yes
Industry-Year FE	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes
Adjusted R^2	0.529	0.530	0.530
Observations	7,903	7,883	7,903
Panel B: Dependent Variable – Log Penalty Amount			
	(1)	(2)	(3)
Post × Share of AI Workers$_{2014}$	81.65*	52.71	28.21
	(37.40)	(50.19)	(32.23)
Post × Share of AI Workers$_{2014}$ × L.Log Assets	-24.38***		
	(6.691)		
Post × Share of AI Workers$_{2014}$ × L.Intangible Capital		-20.21*	
		(8.937)	
Post × Share of AI Workers$_{2014}$ × L.Organizational Capital			-20.19**
			(6.631)
Control Variables	Yes	Yes	Yes
Industry-Year FE	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes
Adjusted R^2	0.483	0.482	0.483
Observations	7,841	7,818	7,840

Note: This table reports Difference-in-Differences (DiD) estimates of the heterogeneous effects of firm-level AI adoption on workplace misconduct. Panel A focuses on the natural logarithm of one plus the number of violations, while Panel B examines the natural logarithm of one plus the penalty amounts. Interaction terms capture heterogeneity by firm size (L.Log Assets), intangible capital (L.Intangible Capital), and organizational capital (L.Organizational Capital). Control variables include a *lag* of Leverage, CashFlow, DividendPayout, FirmSize, Sales, Employees, LaborIntensity, AssetTurnover, Market-to-Book, Tangibility, Capex, R&D Intensity, and FirmAge. Standard errors (in parentheses) are two-way clustered at the firm and year level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 8: DiD Estimates: Other Offense Types

	Log Violations	Log Penalty
Panel A: Competition-related Offenses		
Post × Share of AI Workers$_{2014}$	-0.131	0.769
	(0.705)	(1.642)
Adjusted R^2	0.079	0.035
Observations	5,487	5,474
Panel B: Consumer-protection-related Offenses		
Post × Share of AI Workers$_{2014}$	-0.149	-4.135
	(0.308)	(5.139)
Adjusted R^2	0.537	0.368
Observations	5,484	5,481
Panel C: Environment-related Offenses		
Post × Share of AI Workers$_{2014}$	0.124	0.888
	(0.699)	(6.926)
Adjusted R^2	0.725	0.703
Observations	5,884	5,902
Panel D: Financial Offenses		
Post × Share of AI Workers$_{2014}$	0.130	-0.404
	(0.172)	(2.690)
Adjusted R^2	0.259	0.173
Observations	5,457	5,453
Panel E: Government-contracting-related Offenses		
Post × Share of AI Workers$_{2014}$	-0.232	-4.862
	(0.200)	(2.832)
Adjusted R^2	0.405	0.294
Observations	5,492	5,475
Panel F: Healthcare-related Offenses		
Post × Share of AI Workers$_{2014}$	-0.009	-0.566
	(0.081)	(1.081)
Adjusted R^2	0.822	0.830
Observations	5,436	5,439
Control Variables	Yes	Yes
Industry-Year FE	Yes	Yes
Firm FE	Yes	Yes

Note: This table reports Difference-in-Differences (DiD) estimates of the effect of firm-level AI adoption on workplace misconduct across different offense categories. All regressions include lagged controls for Leverage, Cash Flow, Dividend Payout, FirmSize, Sales, Employment, Labor Intensity, Asset Turnover, Market-to-Book Ratio, Tangibility, Capex, R&D Intensity, and Firm Age. Standard errors are two-way clustered at the firm and year level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 9: DiD Estimates: Alternative Set of Fixed Effects

Panel A: Dependent Variable - Log Violations					
	(1)	(2)	(3)	(4)	(5)
Post × Share of AI Workers$_{2014}$	-4.166***	-5.452***	-5.520***	-7.485***	-7.469***
	(0.646)	(1.197)	(1.204)	(1.591)	(1.606)
Control Variables	No	Yes	Yes	Yes	Yes
Year FE	No	No	Yes	No	Yes
Firm FE	No	No	No	Yes	Yes
Adjusted R^2	0.0380	0.254	0.254	0.519	0.520
Observations	16,689	8,222	8,222	8,179	8,179
Panel B: Dependent Variable - Log Penalty					
	(1)	(2)	(3)	(4)	(5)
Post × Share of AI Workers$_{2014}$	-33.43***	-45.73***	-46.29***	-61.42***	-61.12***
	(5.902)	(8.235)	(8.377)	(10.96)	(10.83)
Control Variables	No	Yes	Yes	Yes	Yes
Year FE	No	No	Yes	No	Yes
Firm FE	No	No	No	Yes	Yes
Adjusted R^2	0.0421	0.243	0.243	0.475	0.476
Observations	16,607	8,154	8,154	8,109	8,109

This table reports Difference-in-Differences (DiD) estimates by incorporating alternative set of fixed effects. Panel A focuses on the natural logarithm of one plus the number of violations, while Panel B examines the natural logarithm of one plus the penalty amounts. Control variables include a *lag* of Leverage, CashFlow, DividendPayout, FirmSize, Sales, Employees, LaborIntensity, AssetTurnover, Market-to-Book, Tangibility, Capex, R&D Intensity, and FirmAge. Standard errors (in parentheses) are two-way clustered at the firm and year level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 10: DiD Estimates: Alternative AI Measure

	(1) Log Violations	(2) Asinh Violations	(3) PPML Violations	(4) Log Penalty	(5) Asinh Penalty	(6) PPML Q(Penalty)
Post × Share of Narrow-AI Workers$_{2014}$	**-4.327***	**-5.633***	**-231.3****	**-45.72****	**-48.97****	**-28.43****
	(2.139)	(2.762)	(92.98)	(20.05)	(21.52)	(13.92)
Leverage	0.0117	0.0155	0.637***	0.123	0.131	0.0887
	(0.00752)	(0.00975)	(0.213)	(0.107)	(0.114)	(0.0703)
CashFlow	-0.00226	-0.00270	-0.353	0.0211	0.0223	0.0387
	(0.00858)	(0.0111)	(0.624)	(0.0884)	(0.0949)	(0.0584)
DividendPayout	0.229	0.294	1.203	2.920	3.145	1.248
	(0.219)	(0.283)	(1.862)	(2.576)	(2.756)	(0.862)
Size	0.0130**	0.0168**	0.189*	0.137**	0.146**	0.0621**
	(0.00542)	(0.00702)	(0.103)	(0.0445)	(0.0479)	(0.0275)
Sales	-0.000770	-0.00103	-0.00159	-0.00367	-0.00370	0.00918
	(0.00173)	(0.00225)	(0.129)	(0.0177)	(0.0190)	(0.0143)
Employment	0.000531	0.000688	-0.000185	-0.00257	-0.00288	-0.00230*
	(0.000806)	(0.00106)	(0.00312)	(0.00494)	(0.00514)	(0.00127)
LaborIntensity	-0.0000936	-0.000140	0.00329	0.00291	0.00305	0.00294
	(0.00105)	(0.00136)	(0.0128)	(0.0134)	(0.0142)	(0.00594)
AssetTurnover	-0.00645	-0.00829	-0.0539	-0.0624	-0.0675	-0.0375
	(0.00880)	(0.0115)	(0.109)	(0.0599)	(0.0645)	(0.0337)
Market-to-Book	-0.000511	-0.000671	0.0234	0.00330	0.00370	0.00439
	(0.00162)	(0.00211)	(0.0498)	(0.0136)	(0.0146)	(0.00906)
Tangibility	-0.00941	-0.0105	0.142	-0.299	-0.323	-0.104
	(0.0642)	(0.0850)	(0.621)	(0.522)	(0.556)	(0.179)
Capex	-0.00751	-0.00662	0.776	-0.431	-0.461	-0.217
	(0.0680)	(0.0885)	(1.095)	(0.489)	(0.520)	(0.263)
R&D Intensity	-0.00118	-0.00173	-5.443**	0.0386	0.0395	-0.0121
	(0.0111)	(0.0144)	(2.722)	(0.109)	(0.117)	(0.0672)
FirmAge	-0.0319	-0.0408	0.178	-0.608*	-0.653*	-0.115
	(0.0331)	(0.0432)	(0.218)	(0.310)	(0.331)	(0.110)
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes
Industry-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.526	0.526	–	0.480	0.479	–
Observations	7,952	7,952	2,826	7,885	7,885	7,885

Note: This table reports Difference-in-Differences (DiD) estimates of the effect of firm-level AI adoption on workplace misconduct using the alternative AI measure of a "narrow" definition of AI skills. Columns (1) and (4) use the $ln(x+1)$ transformation of the dependent variable, columns (2) and (5) use the inverse hyperbolic sine ($asinh(x)$) transformation, and columns (3) and (6) employ Poisson pseudo–maximum likelihood with high-dimensional fixed effects. For column (6), we construct a categorical variable based on the deciles of the total penalty value within each industry (2-digit *NAICS*) and year. Standard errors (in parentheses) are two-way clustered at the firm and year level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 11: DiD Estimates: Additional Control Variables

	(1) Log Violations	(2) Asinh Violations	(3) PPML Violations	(4) Log Penalty	(5) Asinh Penalty	(6) PPML Q(Penalty)
Post × Share of AI Workers$_{2014}$	**-10.33*****	**-13.33*****	**-118.9*****	**-82.69*****	**-88.82*****	**-38.80*****
	(2.283)	(2.945)	(36.10)	(23.25)	(24.90)	(10.76)
Leverage	0.0602	0.0792	0.456**	0.600	0.639	0.251
	(0.0387)	(0.0502)	(0.226)	(0.418)	(0.449)	(0.190)
CashFlow	-0.00480	-0.00476	-0.510	0.234	0.250	0.109
	(0.0720)	(0.0931)	(0.813)	(0.595)	(0.640)	(0.343)
DividendPayout	0.184	0.222	0.847	5.829	6.229	2.113
	(0.538)	(0.692)	(3.414)	(5.882)	(6.308)	(1.579)
Size	0.0302	0.0399	0.298*	0.210	0.221	0.0768
	(0.0227)	(0.0297)	(0.163)	(0.174)	(0.185)	(0.0560)
Sales	0.00779	0.0101	0.138	0.0513	0.0547	0.0330
	(0.00540)	(0.00697)	(0.145)	(0.0573)	(0.0610)	(0.0310)
Employment	-0.00114	-0.00151	-0.00728*	-0.00849	-0.00928	-0.00412**
	(0.00115)	(0.00152)	(0.00423)	(0.00927)	(0.00979)	(0.00173)
LaborIntensity	0.00157	0.00203	0.00588	0.00618	0.00660	0.00167
	(0.00302)	(0.00393)	(0.0166)	(0.0368)	(0.0392)	(0.0106)
AssetTurnover	-0.0212	-0.0280	-0.0684	-0.162	-0.174	-0.0431
	(0.0360)	(0.0471)	(0.212)	(0.268)	(0.285)	(0.0802)
Market-to-Book	-0.00185	-0.00265	-0.000606	0.00656	0.00764	0.00772
	(0.00506)	(0.00660)	(0.0565)	(0.0462)	(0.0496)	(0.0226)
Tangibility	-0.190	-0.246	-0.821	-1.996	-2.142	-0.735**
	(0.168)	(0.223)	(0.805)	(1.246)	(1.328)	(0.304)
Capex	0.0739	0.105	2.564	-1.484	-1.614	-0.673
	(0.209)	(0.271)	(1.765)	(1.412)	(1.507)	(0.608)
R&D Intensity	-0.0355	-0.0423	-4.895**	-0.261	-0.293	-0.176
	(0.0764)	(0.100)	(2.119)	(0.578)	(0.620)	(0.308)
FirmAge	-0.0583	-0.0750	0.326	-0.689	-0.741	0.000595
	(0.0864)	(0.112)	(0.621)	(0.766)	(0.820)	(0.295)
Number of Directors	0.0249	0.0335	0.233	-0.104	-0.107	0.0478
	(0.104)	(0.134)	(0.609)	(1.068)	(1.142)	(0.330)
Non-Executive Director Ratio	0.0205	0.0411	0.302	0.825	0.971	1.137
	(0.844)	(1.094)	(6.381)	(8.475)	(9.091)	(3.034)
Log Delta	0.0280	0.0367	0.149	-0.184	-0.207	-0.0926
	(0.0755)	(0.0980)	(0.373)	(0.671)	(0.719)	(0.205)
Log Vega	0.00591	0.00761	0.0317	0.109	0.117	0.0368
	(0.0119)	(0.0156)	(0.0553)	(0.0804)	(0.0864)	(0.0254)
Log FirmRelatedWealth	-0.0434	-0.0566	-0.264	0.0750	0.0906	0.0565
	(0.0654)	(0.0849)	(0.332)	(0.600)	(0.643)	(0.187)
Log CEOtenure	0.00701	0.00887	0.0459	0.0949	0.102	0.0294
	(0.0108)	(0.0142)	(0.0737)	(0.0980)	(0.105)	(0.0313)
Log CEOage	0.0486	0.0609	-0.0110	0.354	0.364	0.0209
	(0.132)	(0.170)	(0.591)	(1.266)	(1.353)	(0.387)
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes
Industry-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.515	0.516	–	0.458	0.457	–
Observations	3,493	3,493	1,834	3,441	3,441	3,490

Note: This table replicates the benchmark DiD specification while adding further controls (board/governance and compensation variables shown). Columns (1) and (4) use $ln(x+1)$, columns (2) and (5) use $asinh(x)$, and columns (3) and (6) employ PPML with high-dimensional fixed effects. For column (6), we construct a categorical variable based on the deciles of the total penalty value within each industry (2-digit *NAICS*) and year. Standard errors (in parentheses) are two-way clustered at the firm and year level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 12: DiD Estimates: Additional Industries

	(1) Log Violations	(2) Asinh Violations	(3) PPML Violations	(4) Log Penalty	(5) Asinh Penalty	(6) PPML Q(Penalty)
Post × Share of AI Workers$_{2014}$	**-8.918*****	**-11.55*****	**-89.72*****	**-91.98*****	**-98.62*****	**-41.18*****
	(2.253)	(2.899)	(31.09)	(20.40)	(21.87)	(9.821)
Leverage	0.0495	0.0657	0.391*	0.528	0.561	0.211
	(0.0396)	(0.0514)	(0.213)	(0.402)	(0.433)	(0.184)
CashFlow	-0.00550	-0.00598	-0.0528	0.000120	0.000455	-0.00399
	(0.0743)	(0.0963)	(1.022)	(0.694)	(0.745)	(0.399)
DividendPayout	-0.0470	-0.0792	-0.816	2.732	2.969	1.008
	(0.460)	(0.590)	(2.509)	(5.611)	(6.026)	(1.499)
Size	0.0275	0.0364	0.325**	0.248	0.261	0.0899
	(0.0197)	(0.0258)	(0.145)	(0.183)	(0.195)	(0.0601)
Sales	0.00682	0.00871	0.126	0.100	0.107	0.0558
	(0.00799)	(0.0103)	(0.142)	(0.0864)	(0.0919)	(0.0400)
Employment	-0.00122	-0.00157	-0.00812**	-0.00843	-0.00914	-0.00420*
	(0.00125)	(0.00165)	(0.00353)	(0.0123)	(0.0130)	(0.00249)
LaborIntensity	0.00129	0.00162	0.00566	0.00683	0.00692	0.00173
	(0.00368)	(0.00479)	(0.0181)	(0.0415)	(0.0442)	(0.0113)
AssetTurnover	-0.0167	-0.0222	-0.0532	-0.114	-0.123	-0.0377
	(0.0345)	(0.0451)	(0.198)	(0.295)	(0.314)	(0.0907)
Market-to-Book	-0.00172	-0.00247	-0.00276	-0.00481	-0.00431	0.00506
	(0.00487)	(0.00635)	(0.0646)	(0.0402)	(0.0433)	(0.0214)
Tangibility	-0.126	-0.163	-0.432	-1.746	-1.870	-0.653*
	(0.197)	(0.261)	(1.009)	(1.571)	(1.677)	(0.384)
Capex	0.0591	0.0842	1.999	-2.241	-2.417	-0.843
	(0.198)	(0.257)	(1.987)	(1.547)	(1.643)	(0.545)
R&D Intensity	-0.0235	-0.0275	-3.255	-0.318	-0.349	-0.221
	(0.0875)	(0.114)	(2.599)	(0.672)	(0.719)	(0.358)
FirmAge	-0.0966	-0.125	0.0702	-0.724	-0.778	-0.0112
	(0.0846)	(0.110)	(0.663)	(0.822)	(0.879)	(0.308)
Number of Directors	0.0607	0.0807	0.526	0.356	0.380	0.167
	(0.105)	(0.136)	(0.624)	(1.107)	(1.183)	(0.341)
Non-Executive Director Ratio	0.201	0.276	2.858	2.759	2.988	1.521
	(0.824)	(1.069)	(6.420)	(8.821)	(9.460)	(3.317)
Log Delta	0.0233	0.0309	0.105	-0.358	-0.390	-0.150
	(0.0718)	(0.0936)	(0.323)	(0.671)	(0.719)	(0.203)
Log Vega	0.00787	0.0101	0.0569	0.0981	0.106	0.0359
	(0.0120)	(0.0158)	(0.0547)	(0.0886)	(0.0946)	(0.0244)
Log FirmRelatedWealth	-0.0340	-0.0449	-0.221	0.281	0.309	0.124
	(0.0620)	(0.0808)	(0.290)	(0.602)	(0.646)	(0.186)
Log CEOtenure	0.00319	0.00398	0.0139	0.0815	0.0875	0.0287
	(0.00924)	(0.0121)	(0.0668)	(0.0993)	(0.106)	(0.0331)
Log CEOage	0.0435	0.0553	0.192	-0.0388	-0.0523	-0.137
	(0.122)	(0.159)	(0.666)	(1.071)	(1.148)	(0.349)
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes
Industry-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.493	0.494	–	0.453	0.452	–
Observations	3,530	3,530	1,810	3,492	3,492	3,534

Note: This table replicates the benchmark DiD specification while including the industries (Finance, Utilities, and Other) that were excluded from our benchmark sample. Columns (1) and (4) use $ln(x+1)$, columns (2) and (5) use $asinh(x)$, and columns (3) and (6) employ PPML with high-dimensional fixed effects. For column (6), we construct a categorical variable based on the deciles of the total penalty value within each industry (2-digit *NAICS*) and year. Standard errors (in parentheses) are two-way clustered at the firm and year level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 13: DiD Estimates: Propensity Score Matching

	(1) Log Violations	(2) Asinh Violations	(3) Log Penalty	(4) Asinh Penalty
Post × Share of AI Workers$_{2014}$	**-5.658*****	**-7.356*****	**-46.42*****	**-49.75*****
	(1.365)	(1.772)	(13.05)	(13.93)
Leverage	0.00680	0.00897	0.0762	0.0811
	(0.00593)	(0.00767)	(0.0688)	(0.0737)
CashFlow	-0.00158	-0.00178	0.000785	0.000171
	(0.00563)	(0.00728)	(0.0532)	(0.0578)
DividendPayout	0.338	0.430	4.664	5.010
	(0.218)	(0.278)	(2.763)	(2.973)
Size	0.0127**	0.0163**	0.126***	0.135***
	(0.00461)	(0.00599)	(0.0335)	(0.0360)
Sales	-0.000925	-0.00120	-0.0113	-0.0120
	(0.00206)	(0.00265)	(0.0194)	(0.0207)
Employment	0.000908	0.00119	0.00275	0.00276
	(0.00110)	(0.00145)	(0.00655)	(0.00685)
LaborIntensity	-0.000186	-0.000242	-0.00111	-0.00123
	(0.000941)	(0.00122)	(0.0104)	(0.0111)
AssetTurnover	-0.00647	-0.00846	-0.0487	-0.0526
	(0.00802)	(0.0105)	(0.0520)	(0.0558)
Market-to-Book	-0.000554	-0.000706	-0.000616	-0.000690
	(0.00118)	(0.00153)	(0.00860)	(0.00926)
Tangibility	-0.0213	-0.0267	-0.382	-0.408
	(0.0663)	(0.0875)	(0.508)	(0.541)
Capex	0.0322	0.0457	-0.0767	-0.0853
	(0.0605)	(0.0784)	(0.441)	(0.472)
FirmAge	-0.0291	-0.0373	-0.505*	-0.546*
	(0.0222)	(0.0290)	(0.250)	(0.269)
R&D Intensity	0.0126	0.0162	0.128	0.136
	(0.0125)	(0.0163)	(0.127)	(0.136)
Control Variables	Yes	Yes	Yes	Yes
Industry-Year FE	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Propensity Score Matching	Yes	Yes	Yes	Yes
Adjusted R^2	0.545	0.544	0.499	0.499
Observations	7,038	7,038	6,986	6,986

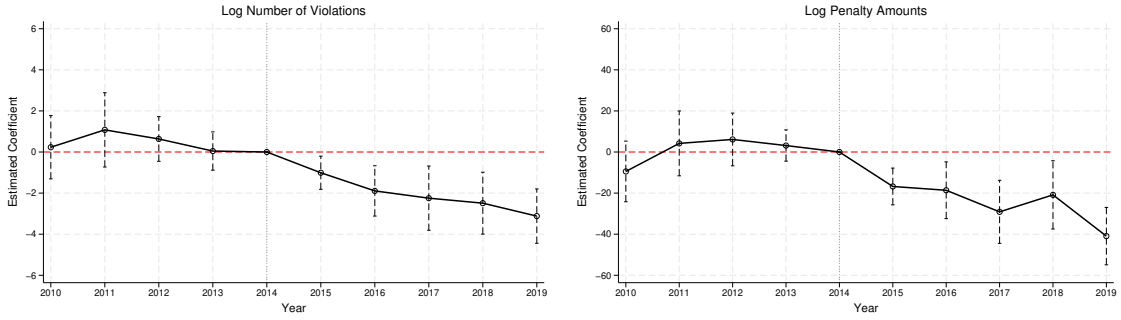
Note: This table reports Difference-in-Differences (DiD) estimates of the effect of firm-level AI adoption on workplace misconduct, implementing a propensity score matching. *Post* equals one for $t \geq 2015$ (Google TensorFlow release). The 2014 share of AI workers proxies for pre-treatment AI intensity. Columns (1) and (3) use the $ln(x+1)$ transformation, while columns (2) and (4) use the inverse hyperbolic sine ($asinh(x)$) transformation. We do not report PPML specifications in this table because the `ppmlhdfe` estimator does not accommodate propensity score matching weights. All specifications include firm, state, and industry-year fixed effects and are estimated on a propensity-score-matched sample. Standard errors (in parentheses) are two-way clustered at the firm and year level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 14: DiD Estimates: Controlling for Other Firm-level Technologies

	(1) Log Violations	(2) Asinh Violations	(3) PPML Violations	(4) Log Penalty	(5) Asinh Penalty	(6) PPML Q(Penalty)
Post × Share of AI Workers$_{2014}$	**-7.737*****	**-10.07*****	**-145.8*****	**-67.71*****	**-72.42*****	**-31.72*****
	(2.042)	(2.652)	(26.71)	(18.39)	(19.58)	(7.509)
Leverage	0.00852	0.0113	0.653***	0.0890	0.0956	0.0832
	(0.0101)	(0.0130)	(0.210)	(0.132)	(0.141)	(0.0790)
CashFlow	0.00279	0.00399	-0.454	0.0597	0.0637	0.0622
	(0.0108)	(0.0138)	(0.644)	(0.115)	(0.124)	(0.0776)
DividendPayout	0.433*	0.555*	2.551	4.567	4.932	1.978**
	(0.228)	(0.296)	(1.770)	(2.799)	(2.994)	(0.797)
Size	0.0119	0.0154	0.1000	0.136*	0.145*	0.0683*
	(0.00660)	(0.00854)	(0.119)	(0.0710)	(0.0764)	(0.0388)
Sales	0.00119	0.00151	0.0218	0.0140	0.0154	0.0206
	(0.00199)	(0.00257)	(0.146)	(0.0202)	(0.0216)	(0.0153)
Employment	-0.000115	-0.000178	-0.00365	-0.00404	-0.00441	-0.00311
	(0.000734)	(0.000948)	(0.00320)	(0.00924)	(0.00977)	(0.00235)
LaborIntensity	-0.000607	-0.000784	-0.0182	-0.00216	-0.00256	-0.000727
	(0.00120)	(0.00157)	(0.0171)	(0.0138)	(0.0147)	(0.00660)
AssetTurnover	-0.00901	-0.0116	-0.0212	-0.0512	-0.0558	-0.0314
	(0.00998)	(0.0130)	(0.111)	(0.0703)	(0.0754)	(0.0327)
Market-to-Book	0.00234	0.00302	0.139***	0.0286	0.0309	0.0191
	(0.00199)	(0.00257)	(0.0522)	(0.0199)	(0.0213)	(0.0119)
Tangibility	0.0157	0.0217	0.473	0.0645	0.0735	0.151
	(0.0641)	(0.0855)	(0.637)	(0.420)	(0.444)	(0.129)
Capex	-0.0247	-0.0271	1.913*	-0.833	-0.902	-0.509**
	(0.0601)	(0.0785)	(1.122)	(0.603)	(0.641)	(0.255)
FirmAge	-0.0318	-0.0398	0.359	-0.871*	-0.932*	-0.104
	(0.0528)	(0.0688)	(0.315)	(0.445)	(0.478)	(0.174)
R&D Intensity	0.00105	0.00122	-5.707***	0.0659	0.0678	0.000266
	(0.0149)	(0.0194)	(1.623)	(0.146)	(0.157)	(0.0856)
Robot Share	0.170	0.148	-14.36	3.911	4.463	2.277
	(1.392)	(1.799)	(16.59)	(11.13)	(12.11)	(5.504)
IT Share	-0.0486	-0.0646	-0.650	0.0438	0.0555	0.112
	(0.0751)	(0.0975)	(1.149)	(0.778)	(0.836)	(0.499)
Data Share	0.194	0.253	-1.820	0.617	0.686	0.355
	(0.222)	(0.288)	(4.950)	(2.556)	(2.747)	(1.586)
Data Analytics Share	-0.00985	-0.0260	3.683	2.366	2.616	1.731
	(0.425)	(0.547)	(6.761)	(4.727)	(5.090)	(2.822)
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes
Industry-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.510	0.510	–	0.464	0.464	–
Observations	6,540	6,540	2,330	6,489	6,489	6,563

Note: This table reports Difference-in-Differences (DiD) estimates of the effect of firm-level AI adoption on workplace misconduct, by controlling for other firm-level technologies provided by Babina et al. (2024). Since these additional controls are available until 2018 in the data, our final period in these DiDs is 2018. *Post* equals one for $t \geq 2015$ (Google TensorFlow release). The 2014 share of AI workers proxies for pre-treatment AI intensity. Columns (1) and (4) use the $ln(x+1)$ transformation; columns (2) and (5) use the inverse hyperbolic sine ($asinh(x)$). Column (3) reports mean violation counts, while column (6) uses deciles of total penalties by industry and year. Standard errors (in parentheses) are two-way clustered at the firm and year level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure 1: Dynamic DiD Estimates: Baseline Results

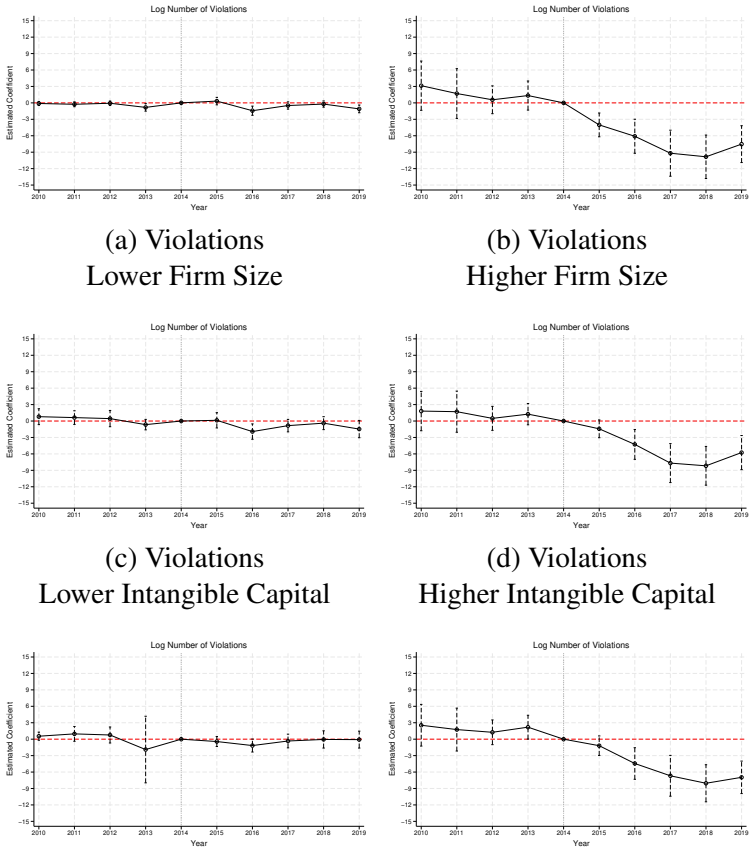


(a) Number of Violations

(b) Penalty Amounts

Note: This figure presents dynamic difference-in-differences estimates for firm-level measures of workplace misconduct (the natural logarithm of one plus the number of violations and the natural logarithm of one plus the penalty amounts). Control variables include a *lag* of Leverage, CashFlow, DividendPayout, FirmSize, Sales, Employees, LaborIntensity, AssetTurnover, Market-to-Book, Tangibility, Capex, R&D Intensity, and FirmAge. The specification includes industry-by-year, and firm fixed effects. Standard errors are two-way clustered at the firm and year level. Confidence intervals are shown at the 95% level.

Figure 2: Dynamic DiD Estimates for Violations: Heterogeneity Analysis



(a) Violations Lower Firm Size

(b) Violations Higher Firm Size

(c) Violations Lower Intangible Capital

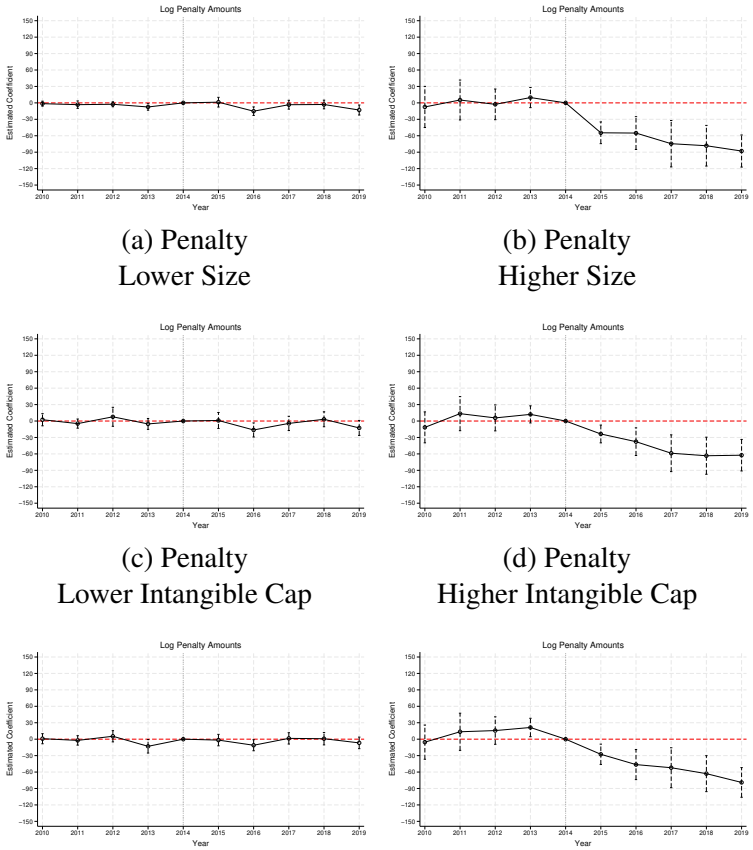
(d) Violations Higher Intangible Capital

(e) Violations Lower Organizational Capital

(f) Violations Higher Organizational Capital

Note: This figure presents dynamic difference-in-differences estimates of the natural logarithm of one plus the number of violations in the heterogeneity analysis. Lower and higher groups are defined as firms below or above the median of the corresponding variable within each year and 2-digit *NAICS* industry. Control variables include a *lag* of Leverage, CashFlow, DividendPayout, FirmSize, Sales, Employees, LaborIntensity, AssetTurnover, Market-to-Book, Tangibility, Capex, R&D Intensity, and FirmAge. The specification includes industry-by-year, and firm fixed effects. Standard errors are two-way clustered at the firm and year level. Confidence intervals are shown at the 95% level.

Figure 3: Dynamic DiD Estimates for Penalty Amounts: Heterogeneity Analysis



(a) Penalty Lower Size

(b) Penalty Higher Size

(c) Penalty Lower Intangible Cap

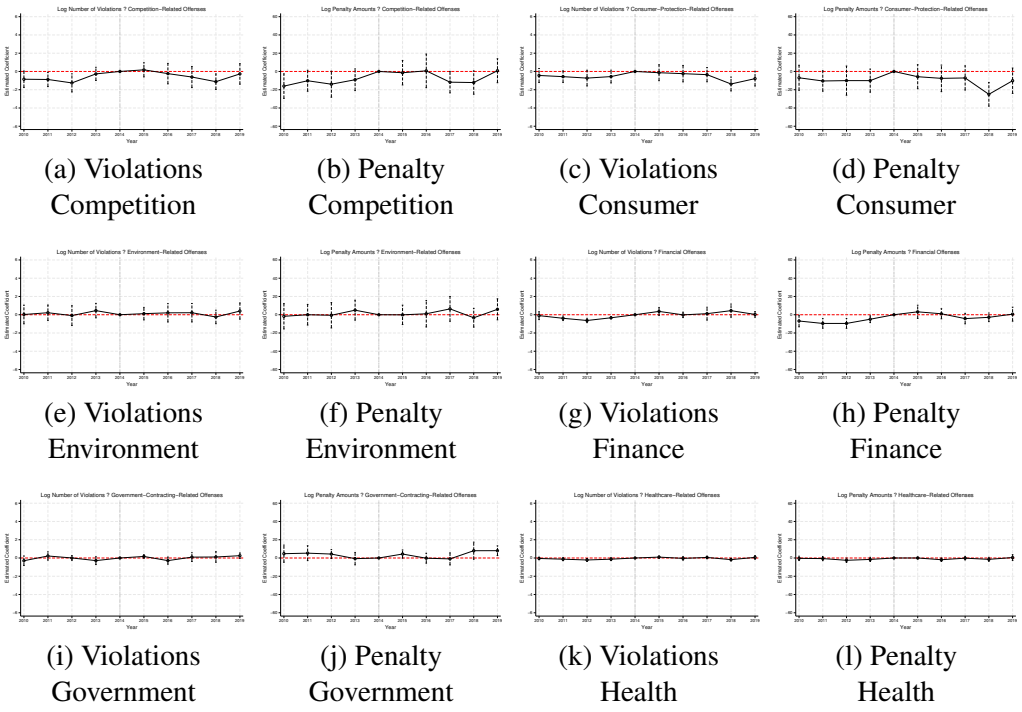
(d) Penalty Higher Intangible Cap

(e) Penalty Lower Organizational Cap

(f) Penalty Higher Organizational Cap

Note: This figure presents dynamic difference-in-differences estimates of the natural logarithm of one plus the penalty amounts in the heterogeneity analysis. Lower and higher groups are defined as firms below or above the median of the corresponding variable within each year and 2-digit *NAICS* industry. Control variables include a lagged of Leverage, CashFlow, DividendPayout, FirmSize, Sales, Employees, LaborIntensity, AssetTurnover, Market-to-Book, Tangibility, Capex, R&D Intensity, and FirmAge. The specification includes industry-by- year, and firm fixed effects. Standard errors are two-way clustered at the firm and year level. Confidence intervals are shown at the 95% level.

Figure 4: Dynamic DiD Estimates: Other Offense Types



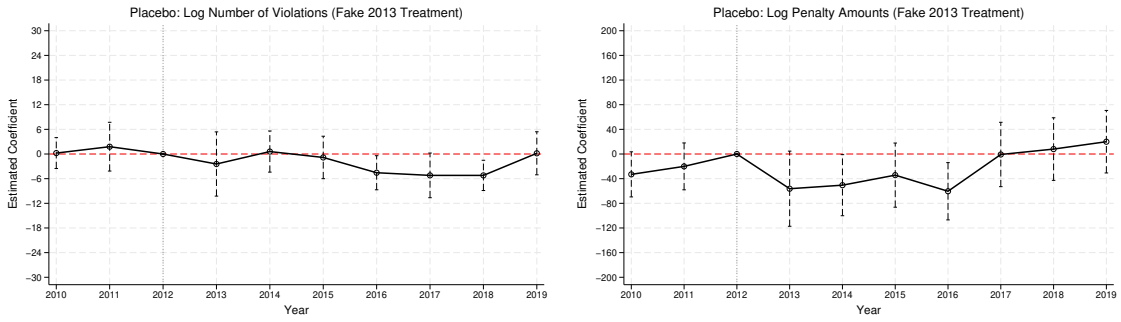
(a) Violations Competition | (b) Penalty Competition | (c) Violations Consumer | (d) Penalty Consumer

(e) Violations Environment | (f) Penalty Environment | (g) Violations Finance | (h) Penalty Finance

(i) Violations Government | (j) Penalty Government | (k) Violations Health | (l) Penalty Health

Note: This figure presents dynamic difference-in-differences estimates for firm-level measures of workplace misconduct (the natural logarithm of one plus the number of violations and one plus the penalty amounts) by offense type. Control variables include a lagged of Leverage, CashFlow, DividendPayout, FirmSize, Sales, Employees, LaborIntensity, AssetTurnover, Market-to-Book, Tangibility, Capex, R&D Intensity, and FirmAge. The specification includes industry-by- year, and firm fixed effects. Standard errors are two-way clustered at the firm and year level. Confidence intervals are shown at the 95% level.

Figure 5: Placebo Test

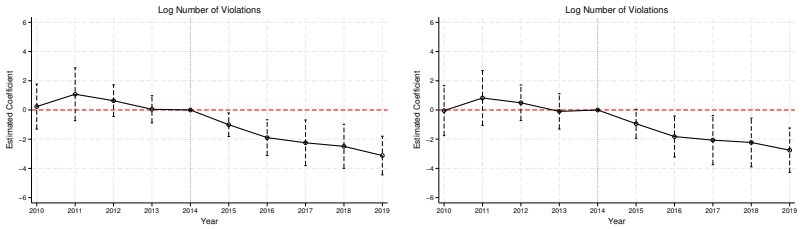


(a) Number of Violations

(b) Penalty Amounts

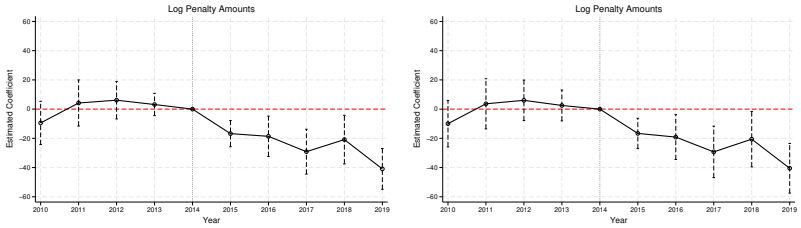
Note: This figure presents dynamic treatment effects for firm-level measures of workplace misconduct (the natural logarithm of one plus the number of violations and one plus the penalty amounts) from a placebo test using year 2014 as the placebo treatment year and 2013 as the baseline. Control variables include a lagged of Leverage, CashFlow, DividendPayout, FirmSize, Sales, Employees, LaborIntensity, AssetTurnover, Market-to-Book, Tangibility, Capex, R&D Intensity, and FirmAge, annualized buy-and-hold returns and annual stock returns, as well as governance variables including Vega, Delta, FirmRelatedWealth, CEOage, CEOtenure, board size, and the ratio of non-executive directors. The specification includes industry-by- year, and firm fixed effects. Standard errors are two-way clustered at the firm and year level. Confidence intervals are shown at the 95% level.

Figure 6: Dynamic DiD Estimates: Alternative Specifications



(a) Violations
FEs + No Trends + Controls

(b) Violations
FEs + Trends + Controls

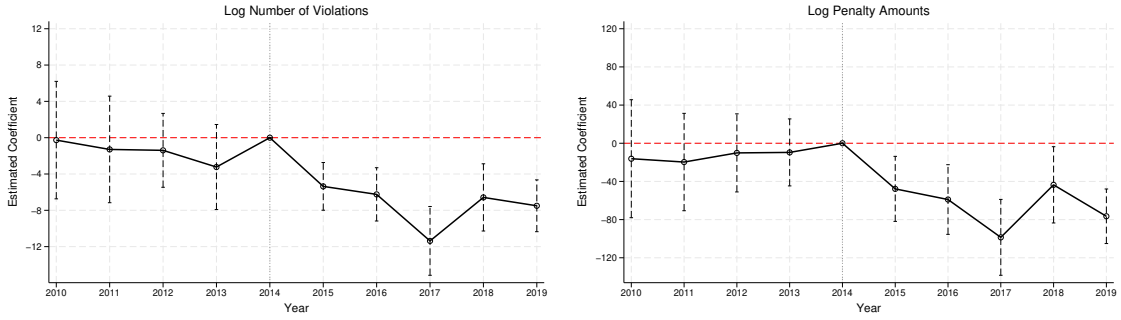


(c) Penalty
FEs + No Trends + Controls

(d) Penalty
FEs + Trends + Controls

Note: This figure presents dynamic difference-in-differences estimates for the firm-level natural logarithm of one plus the number of violations and the natural logarithm of one plus the penalty amounts across various regression specifications. The trends are state-specific and industry-specific linear time trends. Control variables include a lagged of Leverage, CashFlow, DividendPayout, FirmSize, Sales, Employees, LaborIntensity, AssetTurnover, Market-to-Book, Tangibility, Capex, R&D Intensity, and FirmAge, annualized buy-and-hold returns and annual stock returns, as well as governance variables including Vega, Delta, FirmRelatedWealth, CEOage, CEOtenure, board size, and the ratio of non-executive directors. The specification includes industry-by- year, and firm fixed effects. Standard errors are two-way clustered at the firm and year level. Confidence intervals are shown at the 95% level.

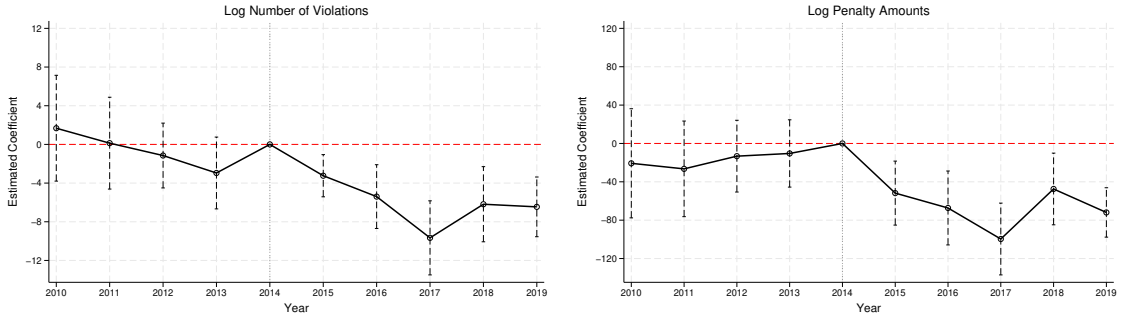
Figure 7: Dynamic DiD Estimates: Additional Control Variables



(a) Number of Violations (b) Penalty Amounts

Note: This table replicates the dynamic difference-in-differences estimates while adding further controls (board/governance and compensation variables shown). Control variables include a lag of Leverage, CashFlow, DividendPayout, FirmSize, Sales, Employees, LaborIntensity, AssetTurnover, Market-to-Book, Tangibility, Capex, R&D Intensity, and FirmAge; as well as governance variables including Vega, Delta, FirmRelatedWealth, CEO age, CEO tenure, board size, and the ratio of non-executive directors. Standard errors are two-way clustered at the firm and year level. Confidence intervals are shown at the 95% level.

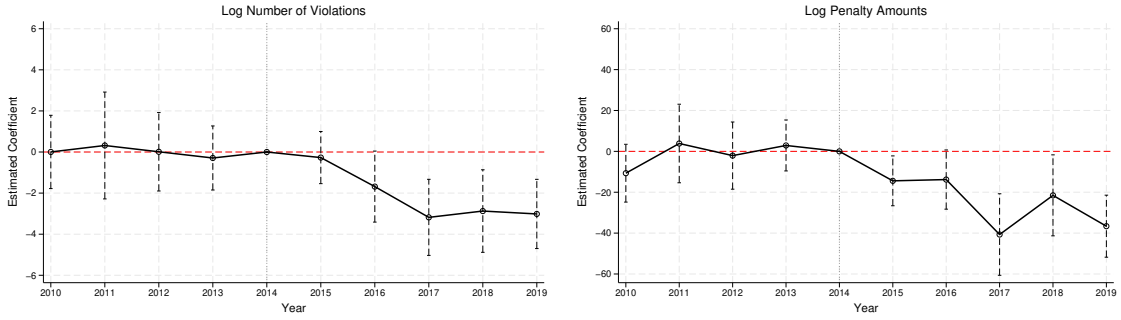
Figure 8: Dynamic DiD Estimates: Additional Industries



(a) Number of Violations (b) Penalty Amounts

Note: This table replicates the dynamic difference-in-differences estimates while including the industries (Finance, Utilities, and Other) that were excluded from our benchmark sample. Control variables include a lag of Leverage, CashFlow, DividendPayout, FirmSize, Sales, Employees, LaborIntensity, AssetTurnover, Market-to-Book, Tangibility, Capex, R&D Intensity, and FirmAge; as well as governance variables including Vega, Delta, FirmRelatedWealth, CEO age, CEO tenure, board size, and the ratio of non-executive directors. Standard errors are two-way clustered at the firm and year level. Confidence intervals are shown at the 95% level.

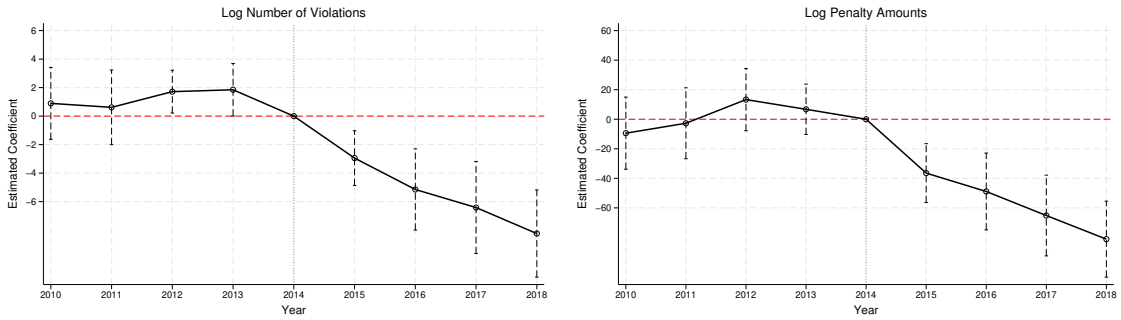
Figure 9: Dynamic DiD Estimates: Propensity Score Matching



(a) Number of Violations (b) Penalty Amounts

Note: This table replicates the dynamic difference-in-differences estimates by implementing a propensity score matching. Control variables include a lag of Leverage, CashFlow, DividendPayout, FirmSize, Sales, Employees, LaborIntensity, AssetTurnover, Market-to-Book, Tangibility, Capex, R&D Intensity, and FirmAge. Standard errors are two-way clustered at the firm and year level. Confidence intervals are shown at the 95% level.

Figure 10: Dynamic DiD Estimates: Controlling for Other Firm-level Technologies



(a) Number of Violations (b) Penalty Amounts

Note: This table replicates the dynamic difference-in-differences estimates by controlling for other firm-level technologies provided by Babina et al. (2024). Since these additional controls are available until 2018 in the data, our final period in these DiDs is 2018. Control variables include a lag of Leverage, CashFlow, DividendPayout, FirmSize, Sales, Employees, LaborIntensity, AssetTurnover, Market-to-Book, Tangibility, Capex, R&D Intensity, and FirmAge. Standard errors are two-way clustered at the firm and year level. Confidence intervals are shown at the 95% level.

Appendix A Additional Tables

Table A.1: Sample Selection: Firm-Year Panel

Step	Dropped Firms	Dropped Observations	Remaining Firms	Remaining Observations
Initial sample (2010-2019)			10,545	66,598
Exclude: SIC 6000-6999	2,559	18,547	7,986	48,051
Exclude: SIC 4000-4999	364	2,861	7,622	45,190
Exclude: SIC $\geq$ 9000	193	800	7,429	44,390
Remaining after industry exclusions			7,429	44,390
After dropping singletons and missing values in regressions	6,322	36,484	1,107	7,906
Final sample			**1,107**	**7,906**

Note: This table reports sample selection for the 2010-2019 firm-year panel with safety and employment-related offense groups. We begin with 10,545 unique firms (66,598 firm-year observations) including all the regression variables with missing and non-missing values. Industry exclusions follow standard practice: financials (SIC 6000–6999), utilities (SIC 4000–4999), and other industries with SIC $\geq$ 9000. The final regression sample excludes singleton and missing values for firm-year panel, leaving 1,107 unique firms and 7,906 firm-year observations between 2010-2019.

Table A.2: Variable Definitions

Variable	Definition	Database
Violations	Total number of safety- and employment-related violations per year	Violation Tracker
Log Violations	Natural logarithm of one plus the number of violations	Violation Tracker
Penalties	Total dollar value of penalties for workplace violations per year	Violation Tracker
Log Penalties	Natural logarithm of one plus the total penalty amount	Violation Tracker
Share of AI Workers	Share of employees with AI-related skills	Babina et al. (2024)
Leverage	Book value of debt ($ million) divided by total assets ($ million)	Compustat
CashFlow	Operating cash flows ($ million) scaled by total assets ($ million)	Compustat
DividendPayout	Cash dividends to common shareholders ($ million) divided by lagged book assets ($ million)	Compustat
Size	Natural logarithm of market capitalization ($ million)	Compustat
Sales	Annual growth in firm sales ($ million)	Compustat
Employees	Natural logarithm of the number of employees (thousand)	Compustat
LaborIntensity	Ratio of employees (thousand) to total assets (normalized to $ thousand)	Compustat
AssetTurnover	Sales ($ million) divided by total assets ($ million)	Compustat
Market-to-Book	Ratio of market ($ million) to book value of assets ($ million)	Compustat
Tangibility	Tangible fixed assets ($ million) divided by total assets ($ million)	Compustat
Capex	Capital expenditures (excluding acquisitions) ($ million) divided by lagged assets ($ million)	Compustat
R&D Intensity	R&D expenditures ($ million) divided by total assets ($ million)	Compustat
FirmAge	Difference between current year and firm's founding year	Compustat
Vega	Dollar change in median executive's equity portfolio value for a 0.01 change in return volatility	BoardEx, Compustat, CRSP, Coles et al. (2006)
Delta	Dollar change in median executive's equity portfolio value for a 1% change in stock price	BoardEx, Compustat, CRSP, Coles et al. (2006)
FirmRelatedWealth	Value of median executive's stock and option portfolio	BoardEx, Compustat, CRSP
CEOage	Age of the CEO (years)	ExecuComp, Coles et al. (2006)
CEOtenure	Number of years since CEO's appointment	ExecuComp
Board Size	Total number of directors on the board	BoardEx
Non-Executive Director Ratio	Ratio of non-executive directors to total board size	BoardEx
Intangible Capital	Firm-level stock of intangible assets constructed using perpetual inventory method ($ million)	Peters and Taylor (2017)
Organizational Capital	Component of SG&A expenses reflecting investment in organizational structures and routines ($ million)	Peters and Taylor (2017)

Notes: This table provides the variables used in our analysis, their definitions, units, and the corresponding data sources.

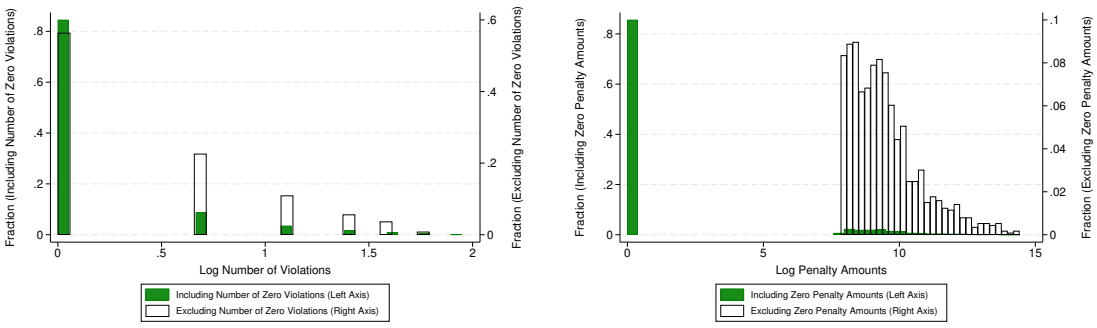
Table A.3: Pairwise Correlation Matrix

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)
(1) Log Violations	1.000																									
(2) Log Penalty	0.947*	1.000																								
(3) Leverage	0.106*	0.097*	1.000																							
(4) CashFlow	0.112*	0.115*	-0.474*	1.000																						
(5) DividendPayout	0.158*	0.152*	0.004	0.175*	1.000																					
(6) Size	0.326*	0.316*	-0.038*	0.344*	0.319*	1.000																				
(7) Sales	-0.048*	-0.047*	-0.006	-0.031*	-0.055*	0.026*	1.000																			
(8) Employment	0.430*	0.390*	0.146*	0.142*	0.169*	0.445*	-0.061*	1.000																		
(9) LaborIntensity	0.057*	0.052*	0.090*	0.034*	-0.005	-0.187*	-0.070*	0.236*	1.000																	
(10) AssetTurnover	0.099*	0.097*	0.062*	0.225*	0.076*	-0.086*	-0.006	0.146*	0.471*	1.000																
(11) Market-to-Book	-0.097*	-0.095*	0.132*	-0.244*	0.066*	0.197*	0.158*	-0.100*	-0.071*	-0.070*	1.000															
(12) Tangibility	0.218*	0.217*	0.093*	0.129*	0.067*	0.051*	-0.086*	0.208*	0.369*	0.210*	-0.143*	1.000														
(13) Capex	0.053*	0.058*	-0.031*	0.116*	-0.008	0.060*	0.033*	0.079*	0.297*	0.258*	0.020*	0.582*	1.000													
(14) R&D Intensity	-0.208*	-0.211*	0.169*	-0.608*	-0.180*	-0.284*	0.099*	-0.220*	-0.149*	-0.266*	0.329*	-0.274*	-0.157*	1.000												
(15) FirmAge	0.226*	0.231*	0.060*	0.210*	0.248*	0.298*	-0.143*	0.297*	0.068*	0.163*	-0.121*	0.137*	-0.031*	-0.276*	1.000											
(16) Number of Directors	0.301*	0.290*	0.213*	0.186*	0.206*	0.595*	-0.056*	0.347*	-0.111*	-0.048*	-0.049*	0.086*	0.004	-0.222*	0.295*	1.000										
(17) Non-Executive Director Ratio	-0.251*	-0.245*	-0.173*	-0.139*	-0.162*	-0.448*	0.031	-0.276*	0.052*	-0.053*	0.032*	-0.153*	-0.071*	0.214*	-0.176*	-0.747*	1.000									
(18) Log Delta	0.166*	0.158*	0.101*	0.263*	0.134*	0.654*	0.080*	0.270*	-0.138*	-0.072*	0.293*	-0.000	0.015	-0.127*	0.075*	0.271*	-0.303*	1.000								
(19) Log Vega	0.117*	0.115*	0.062*	0.092*	0.099*	0.393*	0.034*	0.181*	-0.146*	-0.131*	0.111*	0.011	-0.002	-0.008	0.048*	0.230*	-0.188*	0.635*	1.000							
(20) Log FirmRelatedWealth	0.157*	0.150*	0.104*	0.259*	0.105*	0.632*	0.094*	0.260*	-0.123*	-0.052*	0.307*	0.002	0.027*	-0.125*	0.063*	0.251*	-0.295*	0.985*	0.530*	1.000						
(21) Log CEOtenure	-0.035*	-0.036*	-0.156*	0.012	-0.038*	-0.057*	0.044*	-0.027*	0.019	0.012	0.082*	-0.035*	0.019	0.069*	-0.077*	-0.172*	0.060*	0.140*	0.028*	0.156*	1.000					
(22) Log CEOage	0.141*	0.122*	0.106*	0.012	0.128*	0.134*	-0.028*	0.133*	-0.013	0.007	-0.055*	0.054*	-0.079*	-0.059*	0.216*	0.166*	-0.124*	0.112*	0.018	0.117*	0.126*	1.000				
(23) Log Labor Productivity	-0.369*	-0.362*	-0.103*	-0.337*	-0.269*	-0.454*	0.526*	-0.566*	-0.303*	-0.219*	0.262*	-0.314*	-0.069*	0.445*	-0.457*	-0.396*	0.305*	-0.187*	-0.103*	-0.160*	0.111*	-0.055*	1.000			
(24) Log Total Intangible Capital	0.345*	0.329*	0.106*	0.260*	0.254*	0.855*	-0.047*	0.507*	-0.189*	-0.098*	-0.051*	-0.037*	-0.096*	-0.221*	0.351*	0.627*	-0.445*	0.440*	0.318*	0.412*	-0.145*	0.157*	-0.716*	1.000		
(25) Log Organizational Capital	0.307*	0.307*	0.076*	0.376*	0.287*	0.655*	-0.180*	0.442*	0.040*	0.229*	-0.170*	0.165*	0.064*	-0.496*	0.437*	0.498*	-0.395*	0.283*	0.192*	0.259*	-0.143*	0.081*	-0.727*	0.737*	1.000	
(26) Share of AI Workers	-0.204*	-0.205*	-0.068*	-0.064*	-0.087*	0.004	0.031*	-0.162*	-0.235*	-0.238*	0.116*	-0.297*	-0.107*	0.271*	-0.188*	-0.118*	0.133*	-0.019	-0.034	-0.013	0.014	-0.113*	0.239*	0.002	-0.085*	1.000

Note: This table reports the pairwise correlation matrix. Columns display only index numbers to save space; each row shows the index followed by the variable name. All variables are winsorized at the 1st and 99th percentiles. An asterisk (*) indicates significance at the 10% level.

Appendix B Additional Figures

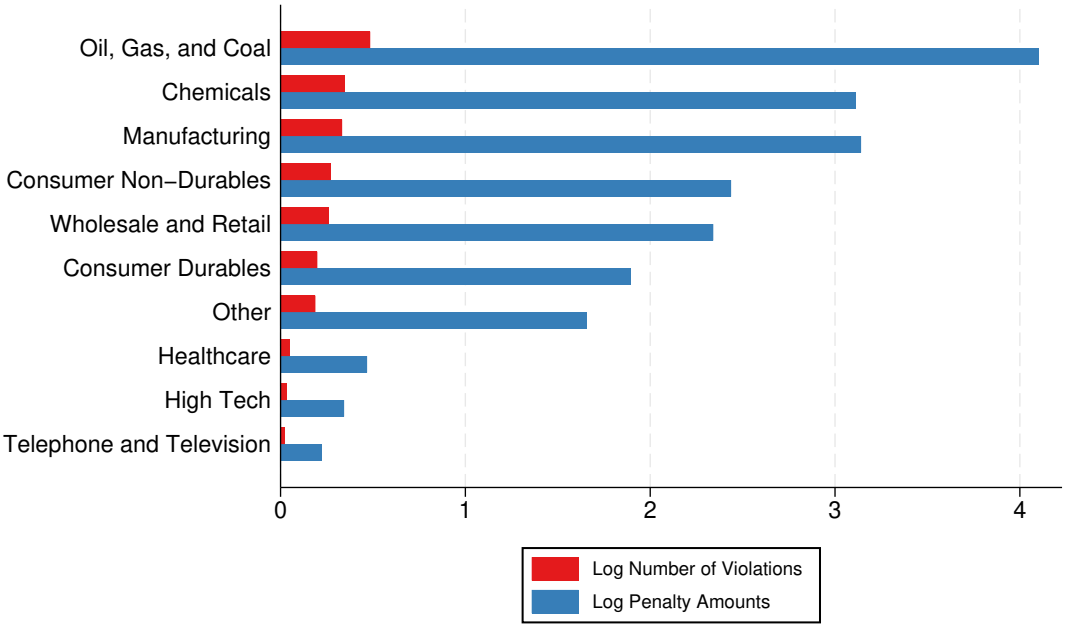
Figure B.1: Histogram - Number of Violations and Penalty Amounts



(a) Number of Violations (b) Penalty Amounts

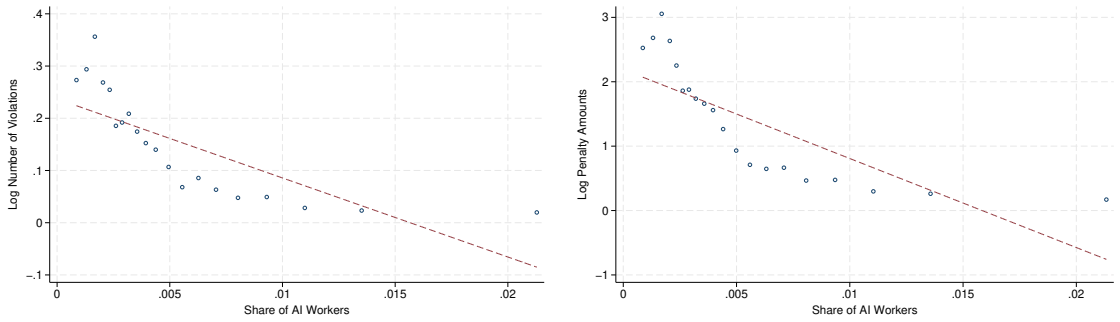
Note: The figure shows histograms of the natural logarithm of one plus the number of violations and the natural logarithm of one plus the penalty amounts, displayed separately for the distribution including zeros (green box) and the distribution excluding zeros (white box).

Figure B.2: Histogram - Number of Violations and Penalty Amounts by Industry



Note: This figure shows the average of the natural logarithm of one plus the number of violations and the natural logarithm of one plus the penalty amounts across Fama-French industries.

Figure B.3: Binscatter Plots: Workplace Misconduct and AI



(a) Number of Violations and AI

(b) Penalty Amounts and AI

Note: This figure presents binned scatterplots of the share of AI workers against the natural logarithm of one plus the number of violations and the natural logarithm of one plus the penalty amounts, respectively.