Auditing Effects on Employment Hiring: Evidence from the New York City Algorithmic Bias Audit Law[1]

Daniel Aobdia (daniel_aobdia@psu.edu)
The Pennsylvania State University

Hao Ma (hao.ma@hhs.se)
Stockholm School of Economics

Sheryl Zhang (zhangs@essec.edu)
ESSEC Business School

This draft: November 2025

Abstract

This paper examines the effects of algorithmic bias audits on workforce diversity by exploiting New York City's (NYC) pioneering Bias Audit Law (Local Law 144). The law requires third-party audits for NYC employers that use algorithm-driven hiring tools to examine instances of recruiting bias. Using a difference-in-differences analysis, we find that bias audit requirements lead to a two percentage-point reduction in the share of male hiring, while having no significant effects on white employee hiring. The effects of decreased male hires are particularly pronounced among firms with greater pre-audit gender imbalances and higher AI intensity. We find that this decline is also concentrated in lower to mid-tier positions, and applies to both public and private companies. We also document costs in the form of longer time to fill vacancies. Overall, our results provide the first large-scale empirical evidence on the effectiveness of mandated algorithmic bias audits, highlighting a trade-off between levelling the playing field for job seekers and reduced recruitment efficiency. They highlight the value of audits beyond financial statements and show that audits can help address AI discrimination in employment practices.

Keywords: Algorithmic bias; Artificial Intelligence; Bias audits; Workforce diversity
JEL Classification Codes: J71; L51; M41; M51; O33

---

[1] Daniel Aobdia was a Senior Economic Research Fellow in the Center of Economic Analysis at the PCAOB between September 2014 and September 2016. He currently advises the PCAOB on questions related to the economics of auditing. The views expressed in this paper are the views of the authors and do not necessarily reflect the views of the board, individual board members, or staff of the PCAOB. We appreciate insightful feedback from Vincent Lin, Henock Louis, Xiumin Martin, Robert Stoumbos, and workshop participants from the Stockholm School of Economics. All errors are our own.

"Automated employment decision tools are in urgent need of transparency, oversight, and regulation. These technologies all too often replicate and amplify bias, discrimination, and harm towards marginalized communities."

Daniel Schwarz, senior privacy and technology strategist at the New York Civil Liberties Union

1. Introduction

In 2024, roughly 80% of U.S. companies and nearly all Fortune 500 companies employed automated systems to screen, score, and rank job market applicants (Dennison 2024). While Artificial Intelligence (AI)-enabled recruitment has the potential to enhance recruitment quality, increase efficiency, and reduce transactional work, it could also pose a significant threat by denying equal access to jobs and discriminating against certain categories of applicants (Chen 2023). Such bias often arises from training data that reflects historical inequalities or underrepresentation, as well as from algorithmic design choices that reinforce discriminatory patterns (Barocas and Selbst 2016; Celiktutan et al. 2024). High-profile lawsuits, such as Derek Mobley's collective action lawsuit alleging that Workday's AI screening disproportionately excluded candidates based on race and age, highlight ongoing concerns about algorithmic bias in hiring processes (Gassam Asare 2025). In response to these concerns, New York City implemented the Bias Audit Law on July 5, 2023. This pioneering legislation requires employers using Automated Employment Decision Tools (AEDTs) to conduct independent annual bias audits and publicly disclose both the use of the AEDTs and the audit results. As a result, bias audit —systematic, independent assurance of internal hiring algorithms for instances of selection disparities —has emerged as a leading policy mechanism to promote fairness and transparency in algorithmic hiring. However, how bias audits affect algorithmic fairness in real-world hiring decisions and the ultimate composition of new hires and the workforce remain open questions. Practitioners, scholars, and policy experts hold sharply divided opinions on the effectiveness of bias audits, underscoring the urgency of evaluating whether bias audits can meaningfully curb algorithmic disparate impacts (e.g., Kroll et al. 2017; Leicht-Deobald et al. 2019; Moschella 2022; Celiktutan et al. 2024; Hilliard et al. 2024; Lam et al. 2024). This paper addresses this critical question by examining the effect of bias audit requirements on hiring diversity.

On the one hand, proponents of bias audits claim that auditing for discrimination is an essential part of the strategy for detecting and responding to biased algorithms. They draw a direct parallel to financial-statement audits, where decades of empirical research show that third-party assurance substantially reduces misstatements (BABL AI 2022; Lam et al. 2024), enhances information quality, and aligns managerial incentives with stakeholder interests (Francis 2004; DeFond and Zhang 2014). By the same logic, systematic audits of hiring algorithms can help diagnose hidden group-based selection gaps. As Kim (2017) asserts, auditing increases firms' awareness of the actual impact of their algorithms on hiring decisions, encouraging corrective actions and deterring firms from engaging in discriminatory hiring practices.

On the other hand, critics argue that bias audits often treat the decision-making process as a black box, merely tracking how the output changes with input, without identifying the root causes of disparities or assessing their substantive significance (Kroll et al. 2017; Leicht-Deobald et al. 2019). This flaw leaves firms without clear remediation guidance and opens the door to dismissing genuine fairness issues or applying superficial tweaks that leave underlying biases intact. Also, employers might simply scale back AI screening in favor of human decision-makers, potentially reintroducing unconscious stereotypes that are far more difficult to detect or remediate at scale (Moschella 2022; Celiktutan et al. 2024).

Despite the controversy, there is little empirical evidence on the effect of bias audits. To fill this gap in the literature, we are the first to provide large-sample empirical evidence to examine the effectiveness of bias audits by investigating New York City's Bias Audit Law (Local Law

144), which is the first law in the U.S to require an AI bias audit. Effective on July 5, 2023, the law prohibits positions located in New York City from using an AEDT in hiring unless the tool has been subject to a public bias audit report within one year of the use of the tool—effectively requiring annual audits for tools in continuous use. These audits must be performed by third parties and check for instances of intentional or unintentional bias built into these systems. At minimum, an independent auditor's evaluation must include calculations of selection or scoring rates and the impact ratio across sex, race/ethnicity, and intersectional categories.

We employ a difference-in-differences analysis to examine the effects of the bias audit on the gender and racial composition of new hires. We are particularly interested in gender because initial experiments using AI recruiting tools, especially a salient one conducted by Amazon in 2014, showed bias against women (e.g., Dastin 2018; Goodman 2018).[2] Using data from Revelio Labs, we construct office-level measures at the year-quarter level. Workforce diversity is measured by the proportion of male hires and white hires.[3,4] Offices in New York City constitute the treated group, while offices outside New York—belonging to the same firms that have at least one office in the city—serve as controls. We then compare how hiring outcomes change differently between treated and control offices around the passage of the Bias Audit Law.

We find that the adoption of bias audits leads to an average two percentage-point drop in the

---

[2] Amazon ultimately cancelled the project. According to Dastin (2018), the bias existed "because Amazon's computer models were trained to vet applicants by observing patterns in resumes submitted to the company over a 10-year period. Most came from men, a reflection of male dominance across the tech industry… In effect, Amazon's system taught itself that male candidates were preferable."

[3] While the Bias Audit Law specifically targets New York City, we use the New York MSA in the New York State as our treated area due to data limitations. This approximation is reasonable given that: (1) New York City comprises approximately 60% of the MSA's total area and contains the majority of commercial districts; (2) the MSA represents a unified labor market with extensive cross-boundary commuting; and (3) most large employers subject to audit requirements maintain primary offices within New York City proper.

[4] One caveat of using the Revelio Labs data is that these data are mostly restricted to professional and white-collar positions. However, using these data also mitigate alternative explanations, such as a concurrent push for more union organizing.

share of male hires, representing a 3.3% reduction of the mean share. We also estimate the dynamic effect of the Bias Audit Law and document little pre-trends, which provides confidence vis-à-vis the parallel trend assumption. An intersectional breakdown shows that this reduction in male hiring occurs uniformly across racial lines: both white-male and non-white-male cohorts experience statistical decreases in their representation. This finding underscores the law's intent to address overlapping dimensions of bias rather than targeting a single demographic group.

In addition, we do not find evidence to suggest that the aggregate share of white hires changes. One possible explanation is that both the Revelio Labs dataset and resume screenings infer race from individuals' names, and name-based algorithms are generally much less accurate in predicting race than gender (Lockhart et al., 2023). Evidence from the Amazon case also indicates that bias against women in AI-based hiring algorithms is common. Additionally, modest shifts within smaller racial/ethnic groups may not visibly alter the aggregate white-versus-non-white ratio. Subsequent granular analysis of the non-white subpopulation reveals statistically significant increases in Black hiring alongside statistically significant declines in Hispanic hiring. These offsetting movements account for the net-zero change in the broad white-hire metric, while still reflecting substantive redistribution of opportunities among non-white groups. Together, these results demonstrate that bias audits can effectuate broad reductions in overrepresentation for some groups and nuanced shifts in racial composition, even when aggregate measures appear unchanged.

To ensure that these results isolate the plausibly causal impact of the Bias Audit Law—and are not driven by New York City's unique culture and initiatives or by other labor-market shifts—we undertake several robustness checks. First, we perform placebo tests on age cohorts: since the law explicitly targets gender and racial categories, it should have no effect on the age composition of hires. As expected, we observe no significant changes in hiring shares across different age groups, confirming that our gender–race findings are not confounded by contemporaneous policies such as pay‑transparency mandates or post‑pandemic labor‑supply shifts. Second, to address concerns that New York City's pronounced commitment to diversity might differentiate it from more typical labor markets, we narrow our control group to offices elsewhere in New York State. Even under this more conservative comparison, we recover the same significant decline in the male‑hire share. Our results also hold under additional robustness checks. Together, these checks bolster the internal validity of our design and strengthen confidence in the law's direct effects on hiring composition.

To shed light on the mechanisms, we explore heterogeneity in the effect of bias audits. Firms are likely to experience greater effects of the Bias Audit Law under two conditions: when they exhibit more severe diversity deficits, as AI tools may inadvertently accentuate such deficits by focusing on characteristics of the more represented group (e.g., Dastin 2018; Goodman 2018), and when they rely more heavily on AI-based hiring screening, as the use of AI is a pre-condition for the Bias Audit Law to be applicable. Consistent with these expectations, we find that firms with the most severe pre-audit gender imbalances and those operating in male-dominated industries experience the largest subsequent reductions in male hiring. Similarly, firms with higher AI intensity—proxied by either prior hiring of AI-skilled workers or operation in high-AI-investment sectors, where algorithmic screening tools are most commonly deployed—demonstrate stronger shifts in gender composition following audit implementation. This alignment between the law's intended focus and its most pronounced effects strengthens our confidence in assessing the impact of mandated audits on improving workforce diversity.

We further investigate which component of the law, third-party auditing or public disclosure, is more likely to drive our main results. Federal contractors provide a useful setting for this analysis: Since 2006, they have been required by the Office of Federal Contract Compliance Programs (OFCCP) to collect and submit applicant demographic data and face OFCCP audit risk to assess equal employment opportunity compliance. However, they are not required to publicly disclose this information, and therefore the New York Bias Audit Law imposes primarily disclosure obligations on them. We find no evidence to suggest a reduction in male hiring among federal contractors, while non-contractor firms—for whom the law introduces both auditing and disclosure requirements—exhibit significant decreases in male hiring. This pattern suggests that the auditing requirement, rather than the disclosure requirement, is the primary mechanism behind the observed effects.

We next conduct two tests to assess the breadth of the law's impact. First, we examine whether the effects are limited to public firms—which face greater regulatory scrutiny—or extend across a broader segment of the labor market. We find that both public and private firms reduce male hiring following the law’s implementation. Second, we test whether the effects vary by career level. The decline in male hiring is particularly pronounced in lower to mid-tier positions, likely because these roles rely more on AI-based employment tools (Godfrey 2025; Jaser et al. 2021) and involve lower adjustment costs in response to fairness mandates (Chang and Kirgios 2024; Larcker et al., 2025).

We also examine potential costs associated with the implementation of the Bias Audit Law. Critics argue that the law may delay hiring processes, prolong staffing shortages, and disrupt overall recruitment operations (Wade 2022). Indeed, mandating third-party audits could create both administrative overhead—commissioning reports and coordinating demographic data—and pressure to extend searches to underrepresented groups to improve diversity performance, potentially reducing recruitment efficiency (Matsa and Miller 2013). Consistent with these concerns, we document a significant increase in the average days to fill a vacancy among treated offices, capturing potential inefficiencies introduced by compliance requirements (Chen and Li 2023). However, we observe no change in the total number of hires, suggesting that firms absorb the delay without curtailing overall recruitment volume or shifting their workforce to offices not subject to the Bias Audit Law.

So far, our results show changes in the composition of new hires, but they might also be explained by companies making arbitrary changes in response to the law that do not necessarily improve fairness in recruiting. To disentangle between both possibilities, we need to investigate whether changes in hiring outcomes following the Bias Audit Law translate into broader firm performance. Improvements in firm performance would be consistent with the law leading to the hiring of more qualified individuals that ultimately lead to improved firm outcomes. This question, however, is difficult to answer because individual office-level performance is not available for most companies, including public and private firms. Instead, we answer a more limited question and focus on financial statement audit firms, where office-level audit quality data are available for their public clients. First, we confirm similar employment patterns for treated audit offices in our sample: these offices reduce the share of male hires in response to the Bias Audit Law. Second, and importantly, we find that treated audit offices' clients exhibit a significant reduction in their fiscal-year-end restatements, a proxy for low audit quality.[5] This evidence thus suggests that audit firms hire more talented individuals, consistent with potential

[5] Restatements can represent a suitable proxy for audit quality because the aim of an audit is to provide assurance that the financial statements are free of misstatements (DeFond and Zhang 2014; Aobdia 2019). We also do not find similar effects for quarterly restatements, consistent with the idea that publicly traded company fiscal year-end financial statements are audited, whereas their quarterly statements are only reviewed. Reviews provide a much more limited level of assurance relative to audits, and auditors do not express an opinion on quarterly financial statements (see PCAOB standard AS 4105).

biases existing prior to the implementation of the law.[6]

This paper makes several contributions to the literature. First, by providing the first large-scale, plausibly causal evidence on the effects of bias audits using a difference-in-differences design, our study offers guidance for policymakers considering similar mandates. We find that third-party audits help rebalance hiring while introducing additional costs in recruitment timelines. These findings can inform the calibration of proposed bias-audit legislation, such as New Jersey's Senate Bill S1588 and Pennsylvania's House Bill HB 594. Beyond immediate policy applications, this research advances the fair AI governance debate by illuminating the practical consequences and tradeoffs that arise when regulating algorithmic hiring technologies to level the playing field and create more equality of chances among different populations.

Second, this study contributes to the emerging literature on AI in corporate decision-making by providing the first evidence of regulatory interventions targeting algorithmic bias in hiring. While recent research has examined AI adoption across various corporate functions—including financial reporting, financial auditing, investment decisions, and capital markets (Bradshaw et al. 2025; Chang et al. 2025; Blankespoor et al. 2025; Choi and Xie 2025; Levy 2025; Munoko et al. 2020)—the intersection of AI governance and human capital management remains largely unexplored. Our work fills this gap by investigating how regulatory oversight can constrain algorithmic bias in hiring, extending the study of AI governance from financial markets to labor markets and contributing to a more comprehensive understanding of how firms can manage AI-related risks across operational domains.

Third, this study extends the scope of third-party auditing beyond financial oversight to

[6] Our results are also consistent with a recent study by Khavis et al. (2025) that shows that an increased proportion of female auditors in an audit office is positively associated with audit quality. In particular, our results suggest that a possible mechanism for their results is that such offices suffer from less hiring bias, and thus hire more qualified individuals for the roles.

employment practices, contributing to an emerging literature on the expanding role of auditors in corporate governance. Building on foundational work on audit independence and quality (Watts and Zimmerman 1983; DeAngelo 1981; DeFond and Zhang 2014; Lawrence et al. 2011), we demonstrate that the monitoring and verification functions central to financial auditing can be effectively applied to algorithmic hiring systems. Our findings complement research on audit scope expansion into non-financial domains, such as sustainability assurance (Aobdia et al. 2025; Gipper et al. 2025) and DeFi assurance (Bourveau et al. 2024; Knechel, Maex, and Park 2025), by showing that employment bias audits generate meaningful behavioral changes in organizational hiring practices.

Finally, this study contributes to the literature on gender diversity. Prior studies have documented the consequences of gender diversity on firms' financial outcomes (e.g., Lins et al., 2024; Liu et al., 2023; Breuer et al., 2024; Billings et al., 2022; Adhikari et al., 2019). Scholars have also explored the mechanisms that promote or hinder progress in corporate diversity, such as same-group representation (Dong 2022; Ahn et al., 2025), diversity quotas or targets (Cai et al., 2024), occupational licensing requirements (Sutherland et al. 2024), disclosure mandate (Bourveau et al., 2025), and diversity information in job postings (Choi et al., 2023). While this literature underscores the importance of institutional design and regulatory levers, much less is known about how emerging technologies—particularly AI-driven hiring tools—affect gender diversity, and whether existing governance mechanisms are sufficient to counteract potential biases. Our study complements this literature by showing that auditing can serve as a mechanism that mitigates algorithm-driven gender bias in hiring, adding to the organizational and regulatory tools that promote fairness in hiring.

2. Institutional Background of New York City's Local Law 144

New York City's Local Law 144 represents the first comprehensive municipal regulation of AI bias in employment decisions in the United States. Effective July 5, 2023, the law prohibits employers and employment agencies from using AEDTs for positions located in New York City, unless these tools have undergone independent bias audits within the previous year—effectively mandating annual audits for any tools in continuous use. Importantly, the law does not require any specific remedial actions based on audit results, focusing instead on increasing transparency. Figure 1 presents the timeline of the New York City Bias Audit Law.

[Insert Figure 1 about here]

An AEDT is defined as a computer-based tool that uses machine learning, statistical modeling, data analytics, or AI to help employers make employment decisions and substantially assist or replace discretionary decision-making. This broad definition encompasses various AI and algorithmic systems used in recruitment, screening, and promotion processes.

The law establishes three core requirements: mandatory independent bias audits conducted annually by third parties, notification to candidates about AEDT usage at least 10 business days prior to evaluation, and public disclosure of audit results on company websites for at least six months. At minimum, bias audits must examine selection rates and impact ratios for gender/sex (male, female, and optionally other) and race/ethnicity categories (Hispanic or Latino, White, Black or African American, Native Hawaiian or Pacific Islander, Asian, Native American), with specific attention to intersectional bias across these dimensions. The selection rates and impact ratios are calculated as follows:

$$selection\,rate = \frac{number\ of\ applicants\ selected\ by\ AEDT\ from\ a\ category}{number\ of\ total\ applicants\ from\ a\ category},$$

and

$$Impact\ ratio = \frac{selection\ rate\ for\ a\ category}{selection\ rate\ of\ the\ most\ selected\ category(eg.\ Male, white\ \ldots.)}.$$

Impact ratios below 0.80 are considered evidence of potential adverse impact under the federal "4/5ths rule," though the law *does not* mandate specific remedial actions when this threshold is crossed. In practice, bias audits are conducted by various types of third parties, including specialized consulting firms (such as Ambrose Consulting, LLC), statistical and economic expert consulting services (such as BLDS, LLC), and algorithmic auditing firms (such as BABL AI Inc.). These auditors typically possess expertise in employment discrimination analysis, statistical modeling, and algorithmic fairness assessment. Bias audits follow a standardized methodology that includes comprehensive system descriptions, methodological frameworks specifying data collection periods and sample sizes, demographic analysis across required categories, statistical calculations of selection and impact ratios, and organizational governance assessments. Appendix B provides a sample audit report.

Enforcement is handled by the New York City Department of Consumer and Worker Protection (DCWP), with civil penalties ranging from $500 to $1,500 per violation and private rights of action available to affected candidates. The law applies to employers using AEDT for positions physically based in the city, regardless of where employers are headquartered. The DCWP clarifies that the key determinant is where the position is physically "located." Bias audits must be conducted for positions located within New York City, whereas roles based elsewhere are exempt. Fully remote positions trigger audit requirements only when the employer's sole offices reside in the city; employers with no NYC offices face no audit obligation, and those operating both inside and outside the city must determine applicability by fact-specific analysis.

Several features make Local Law 144 particularly well-suited for causal identification. The law's geographic specificity creates clear treatment boundaries between New York City and other locations, while its mandatory nature and specific effective date provide clean variation for difference-in-differences analyses. Organizations with operations both inside and outside New York City face identical management and corporate policies, differing only in their exposure to the bias audit requirement.

3. Literature Review and Hypothesis Development

3.1 AI in Corporate Decision-Making and Hiring

The integration of AI in corporate decision-making has fundamentally transformed how organizations operate across multiple domains. Recent literature documents AI's expanding role in financial reporting (Bradshaw et al. 2025), investment decisions (Choi and Xie 2025), and capital markets (Levy 2025), establishing a foundation for understanding AI's broader organizational implications. Within human capital management, this technological shift has been particularly pronounced, with approximately 80% of U.S. companies employing automated systems to screen, score, and rank job applicants in 2024 (Dennison 2024).

There are three key advantages to using AI-driven recruiting. First, sophisticated pattern-recognition algorithms can enhance selection quality by identifying subtle predictors of job performance while minimizing the influence of irrelevant factors and personal biases that affect human screeners (McColl and Michelotti 2019). Second, automating time-intensive processes such as resume parsing and initial screening significantly improves operational efficiency, reducing manual review time and enabling recruiters to focus on higher-value activities (Van den Broek et al. 2021). Third, algorithmic systems standardize hiring procedures by applying consistent, objective criteria across all candidates, theoretically reducing the impact of individual prejudices and increasing fairness in recruitment decisions (Raghavan et al. 2020).

However, this optimistic view confronts mounting anecdotal evidence of algorithmic bias in real-world hiring. In a salient instance, Amazon discontinued a 2014 project to automate hiring because the tool inadvertently became biased against women (Dastin 2018; Goodman 2018). High-profile legal challenges, including Derek Mobley's collective action against Workday alleging that AI screening systems disproportionately excluded candidates based on race and age, also underscore the gap between theoretical potential and practical implementation.

To understand why these high-profile failures occur, the algorithmic bias literature identifies two primary sources of discrimination in AI hiring systems. The first source stems from data quality problems: algorithms produce biased outcomes when trained on inaccurate (Kim 2016), biased (Barocas and Selbst 2016), or unrepresentative input data (Suresh and Guttag 2021). This creates a problematic feedback loop where historical hiring discrimination becomes systematically embedded in algorithmic decisions, perpetuating past inequities rather than correcting them. The second mechanism involves proxy discrimination, where seemingly neutral variables—such as zip codes, educational institutions, or word choices in resumes—serve as statistical proxies for protected characteristics like race or gender (Prince and Schwarcz 2019). Even when algorithms explicitly exclude protected attributes, these correlated variables enable indirect discrimination that may be difficult to detect or prove in legal contexts.

Taken together, these mechanisms suggest that without active intervention, AI systems may actually worsen employment discrimination rather than ameliorate it.

3.2 Policy Interventions in Labor Markets and Employment Discrimination

The economics literature on employment discrimination provides crucial context for understanding regulatory responses to algorithmic bias. Building on foundational work by Becker (1957) and Arrow (1973), scholars have documented persistent labor market disparities across gender and racial lines, motivating decades of policy interventions aimed at promoting equal employment opportunity.

These policy interventions utilize distinct approaches with varying enforcement methods, from direct numerical requirements to transparency-based oversight mechanisms. Norway's 2003 gender quota mandating 40 percent female board representation drove rapid increases in women's corporate board participation and improved market valuations, illustrating how enforceable targets can reshape organizational composition (Ahern and Dittmar, 2012). However, such direct mandates can also trigger unintended displacement effects, as demonstrated by California's SB 826, which produced immediate increases in female directors but also prompted backlash in non-regulated hiring channels (Bian et al. 2025). More subtly, Adams and Ferreira (2009) find that while female directors improve board monitoring practices, the average effect of gender diversity on firm performance is negative, particularly in well-governed firms. This suggests that mandatory diversity requirements can lead to "over-monitoring" that reduces shareholder value, demonstrating how well-intentioned mandates may inadvertently harm organizational performance.

Alternatively, disclosure-based approaches emphasize transparency and public accountability to drive behavioral change. Research on diversity reporting shows that organizations exhibit greater compositional improvements when they publish concrete targets rather than vague commitments (Cai et al. 2024). However, disclosure-only policies face risks of superficial compliance, with some organizations engaging in "diversity washing" to meet reporting obligations without substantively altering workforce demographics (Baker et al. 2024).

New York City's Bias Audit Law combines elements of both approaches, requiring independent third-party evaluation while mandating public disclosure of results. This hybrid regulatory design mirrors successful aspects of earlier interventions: like board gender quotas, it imposes binding compliance requirements, while like diversity disclosure mandates, it emphasizes transparency and public accountability. The law's focus on algorithmic systems presents novel challenges but follows established patterns where regulatory oversight drives compositional changes when audit findings translate into corrective actions rather than cosmetic adjustments, providing the theoretical foundation for our empirical investigation of whether mandatory bias audits can effectively alter hiring practices while avoiding the displacement effects observed in earlier interventions.

3.3 Third-Party Auditing Theory and Effectiveness

The theoretical foundation for third-party auditing rests on information economics and agency theory, which explains how independent verification can mitigate information asymmetries and align incentives between principals and agents. In financial markets, decades of research demonstrate that external audits improve information quality, reduce earnings management, and enhance investor confidence (Francis 2004; DeFond and Zhang, 2014). The monitoring hypothesis suggests that audit scrutiny deters opportunistic behavior by increasing the likelihood of detection and reputational consequences (Watts and Zimmerman, 1983).

The expansion of auditing beyond financial statements has gained momentum in recent years, with assurance services emerging in several domains, including ESG metrics and decentralized protocols. This scope expansion reflects growing stakeholder demand for independent verification of corporate claims across multiple dimensions of performance. Recent studies highlight how ESG assurance is becoming increasingly prevalent in the U.S., with S&P 500 firms' adoption rising from 16% to 46% between 2010-2020 (Gipper et al., 2025). Similarly, the DeFi ecosystem has developed smart contract audits—voluntary code integrity assessments performed by specialized technical firms rather than traditional accounting firms (Bourveau et al., 2024). Building on these developments, our study extends the role of audits into the algorithmic domain, where specialized AI consulting firms and technology companies like BABL AI currently lead the market for algorithmic bias audits, reflecting the technical expertise required to evaluate machine learning systems—similar to how non-traditional audit providers with specialized knowledge dominate ESG and DeFi assurance markets.

The application of auditing principles to algorithmic systems has generated substantial scholarly interest as policymakers seek mechanisms to address AI bias in employment decisions. The theoretical foundation for algorithmic auditing draws directly from established audit theory, where independent third-party verification serves to reduce information asymmetries and align organizational behavior with stakeholder interests (Francis 2004; DeFond and Zhang 2014). This framework suggests that systematic audits of hiring algorithms can help diagnose hidden group-based selection disparities. Kim (2017) argues that auditing increases firms' awareness of algorithmic impacts on hiring decisions, encouraging corrective actions and deterring discriminatory practices. This perspective aligns with broader audit theory, suggesting that third-party oversight improves organizational behavior even when audit methodologies are imperfect.

However, the effectiveness of bias audits in practice remains contested in the literature. A stream of research highlights fundamental limitations in current audit approaches. Kroll et al. (2017) and Leicht-Deobald et al. (2019) demonstrate that algorithmic audits often treat decision processes as black boxes, documenting disparate outcomes without identifying root causes or assessing substantive significance. This approach provides limited remediation guidance and may enable superficial adjustments that preserve underlying biases while creating an appearance of compliance. Additional concerns emerge regarding unintended consequences of audit mandates. Moschella (2022) and Celiktutan et al. (2024) warn that regulatory pressure may prompt employers to abandon AI screening in favor of human decision-makers, potentially reintroducing unconscious stereotypes that are more difficult to detect systematically. This substitution effect could undermine the policy objective of reducing employment discrimination.

The tension between these perspectives reflects broader challenges in AI governance, where technical complexity intersects with legal and social objectives. The implementation of New York City's Bias Audit Law provides an opportunity to empirically evaluate these competing theoretical predictions about algorithmic audit effectiveness in a real-world regulatory context. While prior research has largely relied on theoretical arguments and case studies, the mandatory nature of this regulation and its clear implementation timeline enable plausibly causal identification of bias audit effects on hiring outcomes.

3.4 Hypothesis Development

Drawing from the theoretical frameworks above, we develop hypotheses that address three key empirical questions: First, do mandatory bias audits meaningfully alter hiring composition in ways consistent with reducing algorithmic discrimination? Second, under what conditions are these effects most pronounced? Third, what are the implementation costs of audit compliance?

Hypothesis 1 (Direct Effects on Hiring Composition): We test the null hypothesis that mandatory bias audits have no effect on hiring composition. If bias audits merely treat algorithmic systems as "black boxes" without identifying root causes of discrimination, prompt organizations to abandon AI screening for potentially more biased human decision-making, or generate only cosmetic adjustments that preserve underlying biases, we should observe no systematic changes in hiring patterns following audit implementation. However, if bias audits identify algorithmic discrimination, and organizations effectively remediate the underlying AI issues, we should observe systematic changes in hiring patterns following audit implementation.

The monitoring theory suggests that third-party oversight creates incentives for organizations to modify practices that generate disparate outcomes, even if underlying preferences remain unchanged. We therefore examine whether audit-induced corrections manifest as increased representation of previously underrepresented groups. We are particularly interested in understanding the share of female and minority hires, given anecdotal evidence that AI recruiting tools can mimic prior biases, even unintentionally, for these categories.

H1a: *Bias audits have no effect on the share of male hires.*

H1b: *Bias audits have no effect on the share of minority hires.*

Hypothesis 2 (Heterogeneous Effects by Firm Characteristics): The impact of bias audits should vary systematically based on characteristics that proxy for both the likelihood of algorithmic bias and the scope for remediation. Firms with severe existing imbalances may employ AI recruiting tools that inadvertently exacerbate existing imbalances, because of the way such tools are trained to begin with, while organizations with higher AI intensity are more likely to rely on the algorithmic systems targeted by the law. This heterogeneity allows us to test whether audit effects concentrate in contexts where they are theoretically most needed.

H2a: *Bias audits have stronger effects when more severe pre-existing diversity deficits exist*

H2b: *Bias audits have stronger effects for firms with higher AI intensity*

Hypothesis 3 (Implementation Costs and Efficiency Trade-offs): Compliance with bias audit requirements imposes both direct costs (commissioning audits, coordinating with third parties) and indirect costs (expanding candidate searches, modifying screening processes). These compliance burdens should manifest in longer hiring timelines as organizations navigate new procedural requirements. We propose the following null hypothesis.

H3: *Bias audits have no effect on hiring timelines and total hiring volume.*

These hypotheses collectively address the central research question of whether mandatory third-party audits can effectively reduce algorithmic bias in hiring while identifying the conditions under which such interventions are most impactful and the trade-offs they entail.

4. Research Design and Data

4.1 Research design

To examine the effect of the Bias Audit Law on hiring outcomes (H1a and H1b), we use a difference-in-differences (DiD) approach that exploits geographic variation in the law's implementation. We compare changes in hiring outcomes between treatment offices, located in New York, and control offices, which are offices of the same company but located outside New York.[7] The control sample is restricted to firms that have at least one office in New York to ensure within-firm comparability. This restriction ensures that both treatment and control offices belong to firms that are subject to the same corporate policies, culture, and strategic decisions, with the key difference being their exposure to the Bias Audit Law.

We focus on a nine-quarter window centered on the quarter in which the regulation was enforced (2023Q3). Figure 2 illustrates the research design. Panel A shows the event window timeline, while Panel B depicts the identification of treatment and control offices. Our office-quarter level regression specification is stated in equation (1):

$$Y_{i,t} = \beta_0 + \beta_1 Treat_i * Post_t + \gamma X_{i,t} + u_i + v_t + \varepsilon_{i,t} \tag{1}$$

where *i* indexes offices and *t* quarters. $Y_{i,t}$ is hiring diversity, comprised of two measures: *MaleProp* and *WhiteProp*, the proportions of male and white employees, respectively, hired in office *i* during quarter *t*. These variables allow us to examine how gender and racial

[7] Because Revelio Labs provides location data only at the state and metro area levels, we define the New York area as both the state being New York and the metro area being New York. This is reasonable because New York City comprises 60% of the metro area, functions as a unified labor market with significant commuting, and hosts most large employers subject to audit requirements. Additionally, if the same company appears in different states or metro areas, we treat each as a separate office.

compositions of new hires respond to the regulation.

[Insert Figure 2 about here]

The key explanatory variable is the interaction between *Treat* and *Post*. The indicator variable, *Treat*, equals one for offices located in the New York MSA. The other indicator variable, *Post*, equals one for quarters on or after 2023 Q3, when the Bias Audit Law became effective. The coefficient of interest, β_1, captures the differential change in hiring outcomes for New York offices relative to control offices following the implementation of the Bias Audit Law. We control for *LogEmpNum*, the natural logarithm of the number of employees in office *i* during quarter *t*. To mitigate the influence of outliers, we winsorize all continuous variables at the 1% and 99% levels. All detailed variable definitions are provided in Appendix Table A.

We include office and quarter fixed effects to control for time-invariant characteristics such as long-standing HR practices and to absorb aggregate shocks and seasonal patterns such as nationwide recruitment cycles and hiring surges. Thus, the stand-alone coefficients on *Treat* and *Post* are subsumed by this fixed effect structure. Standard errors are clustered at the office level to account for potential serial correlation in error terms within offices over time.

4.2 Data and sample selection

We construct our sample using data from Revelio Labs, which collects and aggregates all publicly available professional profiles and job postings.[8] The primary data sources for our analysis are the Revelio Individual, Revelio Workforce Dynamics, and Revelio Job Postings databases. The Revelio Individual database provides an extensive range of variables for each

[8] We acknowledge that Revelio Labs primarily captures professional and white-collar workers, which limits the generalizability of our findings to this segment of the labor market rather than to blue-collar or service workers. However, this data characteristic also helps address concerns about confounding events such as strikes and union organizing activities, as professional and white-collar workers are substantially less unionized.

individual, including demographic details such as gender and ethnicity, educational backgrounds such as academic degrees and start and end dates, and employment histories such as location, seniority level, and company industry. The Revelio Workforce Dynamics database includes the total number of employees at each office on a monthly basis. The Revelio Job Postings database contains job postings sourced from aggregator sites and company websites. Our main sample draws from these databases, focusing on US offices from 2022 Q2 to 2024 Q3. For both the Revelio Individual and Job Postings databases, we aggregate the data at the office-quarter level. The Workforce Dynamics data is aggregated quarterly and then merged with the Individual and Job Postings databases for the respective hiring diversity and recruitment efficiency regressions.

We require each office to have both pre- and post-treatment periods. Additionally, each treated office must have a corresponding control office located outside the New York area but within the same company. Our final sample comprises 4,414 companies, 108,856 offices, and 556,551 office-quarter observations for the main hiring regression.[9] Sample sizes vary slightly across different tests due to the use of different dependent variables.

5. Results

5.1 Descriptive statistics

Table 1 reports summary statistics for outcome and control variables. For the full sample, the average proportion of male hires (*MaleProp*) is 57.1%, and the average proportion of white hires (*WhiteProp*) is 70.1%. The table also includes the proportions of different age groups and races. The average log job vacancy duration (*LogDuration*) is 3.551, corresponding to about 35

[9] Our initial sample from the aggregated Revelio Individual database contains 150,727 companies and 645,460 offices. After merging this data with the Workforce Dynamics database to obtain the control variable for the total number of employees, the sample size is significantly reduced. However, our results remain robust when using the large initial sample.

days.[10] The average number of new hires in an office is 5.569. The average log number of employees (*LogEmpNum*) is 4.757, indicating a typical office size of about 117 employees.

These variables exhibit substantial variation across offices and quarters. *MaleProp* has a standard deviation of 0.398, with a 0.125-1.000 interquartile range, indicating that some offices hire predominantly female workers while others hire exclusively male workers. *WhiteProp* also varies considerably, though more moderately than *MaleProp* (SD = 0.365, interquartile range: 0.500-1.000), suggesting greater gender imbalance relative to ethnicity / racial imbalance across offices. The log-transformed variables also display considerable variation, with *LogDuration* ranging from 1.386 to 4.812 and *logEmpNum* from 3.477 to 6.003 across the interquartile range.

[Insert Table 1 about here]

5.2 Bias Audit and Hiring Diversity

We employ the Difference-in-Differences design from Equation (1) to compare hiring outcomes between offices located inside and outside New York, before and after the enforcement of Bias Audit Laws. Table 2 Panel A reports the results, with Columns (1) and (2) for the proportion of male hires, and columns (3) and (4) for the proportion of white hires. For each outcome, we report specifications both without control variables (columns 1 and 3) and with the office size control variable *LogEmpNum* (columns 2 and 4).

The coefficients on the interaction term *Treat*×*Post* in the first two columns of Table 2 Panel A are statistically negative and economically meaningful (−0.019, t = −3.640 in column 1; −0.019, t = −3.660 in column 2). This suggests that New York offices relatively reduce their proportion of male hires on average by almost 2 percentage points following the implementation of the Bias Audit Law, representing a 3.3% reduction from the mean share of male hires. In

[10] The average for the unlogged *Duration* is higher, at 40 days, consistent with some skew existing in the variable.

contrast, the coefficients for *WhiteProp* are insignificant across both specifications. These results suggest an average decline in male hiring after the implementation of the Bias Audit Laws, while no evidence is found for white hiring, thus providing evidence to reject H1a but not H1b.

One possible explanation for this difference is the disparity in algorithmic identification accuracy. Both gender and race information in the Revelio Labs dataset and in resume screenings are inferred from individuals' names. However, name-based algorithms generally predict gender with substantially higher accuracy than race (Lockhart et al. 2023). Moreover, evidence from the Amazon case suggests that bias against women in AI-based hiring algorithms is common.

The coefficient on the control variable *LogEmpNum* is consistently negative and statistically significant, indicating that larger offices tend to have more diverse hiring patterns. This finding aligns with previous literature documenting that larger organizations typically maintain more formalized diversity practices (Oliveira and Zhang 2022).

[Insert Table 2 about here]

A key underlying assumption is that, in the absence of the Bias Audit Laws, the trends for the treatment and control groups would have remained parallel. While this assumption cannot be directly tested, we can still assess whether the pre-treatment trends are parallel. We re-estimate Equation (1) by replacing the *Post* variable with separate indicators for each quarter before and after the regulation. The benchmark quarter is 2023Q2, which immediately precedes enforcement of Bias Audit Regulations. The dynamic results for the male hiring proportions are shown in Figure 3. This figure displays the point estimates and the 90% confidence intervals for each quarter. Point estimates for all quarters before 2023 Q3 are statistically indistinguishable from zero, suggesting no systematic differences between treatment and control offices prior to the law's enforcement. In contrast, we observe a relatively significant decrease in male hiring after the enforcement of the Bias Audit Law.

[Insert Figure 3 about here]

5.3 Bias Audits and Hiring Diversity: White Males and Non-White Males

We next examine how the Bias Audit Law affects male hiring across racial lines. This allows us to determine whether the reduction in male hiring is concentrated among specific groups or represents a broad shift toward greater gender diversity across all racial categories.

The corresponding results are reported in Panel B of Table 2. Column (1) shows the proportion of white-male hires, and column (2) shows the proportion of non-white male hires. The coefficients on *Treat*×*Post* are significantly negative in both columns (−0.009, t = −1.698 for white males; −0.010, t = −2.405 for non-white males), indicating statistically significant decreases in representation for both cohorts. This finding is consistent with the law's broader aim of addressing multiple dimensions of bias, rather than targeting a single demographic group.

5.4 Addressing Confounding Events with Placebo Tests: Bias Audits and Age Diversity in Hiring

A concern is that our results may reflect broader contemporaneous changes in New York City's labor market, such as the city's strong diversity culture and initiatives or other labor-market shifts, rather than the specific impact of bias audits. To strengthen the causal interpretation of our findings and rule out alternative explanations, we conduct placebo tests examining age diversity in hiring patterns. The Bias Audit Law requires audits for race and gender categories but does not specifically require addressing age-based discrimination (e.g., Hickock 2022). This institutional feature provides a natural falsification test: if our observed effects on gender hiring truly result from the Bias Audit Law rather than confounding factors, we should observe no systematic changes in age-related hiring patterns following the law.

We construct age-based hiring proportions using five distinct cohorts based on standard demographic classifications employed by the U.S. Bureau of Labor Statistics: employees under 25 (*PropAgeUnder25*), ages 25-34 (*PropAge25To34*), ages 35-44 (*PropAge35To44*), ages 45-54 (*PropAge45To54*), and ages 55 and above (*PropAge55Plus*). The placebo test results, presented in Table 3, show no statistically significant changes in hiring shares across any age group following implementation of the Bias Audit Law. This finding supports our conclusion that the observed changes in gender diversity are not confounded by other contemporaneous policies, such as diversity campaigns, pay transparency mandates, shifts in labor supply following the pandemic, or the candidate notification requirement under the law.[11]

[Insert Table 3 about here]

6. Additional analyses

Having established that the Bias Audit Law increases female hiring, we explore plausible underlying mechanisms and boundary conditions of these effects. We examine heterogeneity across organizational characteristics (pre-law gender imbalances, H2a, and AI intensity, H2b), firm types (federal versus non-federal contractors and private versus public firms), job levels, and demographic groups. We also assess the potential costs and benefits of the Bias Audit Law by analyzing its impact on recruitment efficiency and firm performance following its enforcement, and conduct robustness tests to validate our main findings.

6.1. Cross-Sectional Tests: Pre-Law Male Hiring Share

Firms with more severe diversity deficits are more likely to be affected by the Bias Audit Law, because overrepresentation of some groups may lead AI-based screening models to put too

[11] If candidate notifications about the use of AEDT in hiring were the major driver of our results, we would expect shifts in age composition, since age discrimination is documented in AEDT use, such as iTutorGroup's lawsuit for using AI to exclude older applicants (Wiessner 2023). The absence of such changes supports attributing the observed effects to the bias audit itself.

much emphasis on the characteristics of such groups, eventually leading to inadvertent screening out of qualified candidates from other groups (H2a). This is what happened when Amazon attempted to implement an AI-based hiring tool in 2014 (Dastin 2018; Goodman 2018). We thus expect the reduction in male hiring to be more pronounced in firms with more severe pre-audit gender imbalances. To test this prediction, we construct (1) *HighMaleFirm*, a measure of pre-law male hiring imbalance which equals one if a firm's pre-law male hire share is above the sample median, and (2) *HighMaleInd*, a measure of industry-level gender composition which equals one if a firm is in a male-dominated industry, defined as industries where the female employment share is below the overall average across all industries. We then estimate triple-difference specifications by interacting these indicators with the *Treat*×*Post* and other indicator variables.

Table 4 presents consistent results with our expectations. In columns (1) and (2), the triple interaction terms are significantly negative, suggesting that companies with more severe pre-audit gender imbalances or in industries with large gender disparities experience larger reductions in male hiring (H2a).

[Insert Table 4 about here]

6.2. Cross-Sectional Tests: AI Intensity

We next examine another potential mechanism: firms' reliance on AI-based hiring screening. Because the Bias Audit laws target algorithmic discrimination, firms that depend more heavily on AI for screening applicants are more likely to be affected. This is particularly relevant for firms with greater AI intensity, where the use of algorithmic hiring practices is widespread. We expect firms that have invested heavily in AI capabilities or operate in AI-intensive industries to exhibit larger reductions in male hiring following the law's enforcement (H2b).

To test this prediction, we construct two binary indicators. *HighAIFirm*, equals one if the office belongs to a top AI-investing firm, proxied by whether it hired AI-skilled workers before the Bias Audit Law, and zero otherwise, following Babina et al. (2024). *HighAIInd* equals one if the company's industry is among the top AI-investing sectors, based on Table 2 in Bonfiglioli et al. (2025). We then estimate the triple-difference specification by interacting each *HighAI* measure with the *Treat*×*Post* indicator. Table 5 presents the results. Confirming our predictions, the triple interaction term is significantly negative in columns (1) and (2). Overall, these results provide evidence that offices with higher AI intensity experience greater reductions in male hiring following the implementation of Bias Audit laws (H2b).

[Insert Table 5 about here]

6.3 Subsample Analyses: Federal and Nonfederal contractors

The Bias Audit Law mandates both an independent audit and the public disclosure of audit results. In this subsection, we explore which component, auditing or disclosure, is more likely to drive the observed changes in hiring diversity.

Federal contractors provide a useful setting for isolating these mechanisms. Since February 6, 2006, the Office of Federal Contract Compliance Programs (OFCCP) has required federal contractors and subcontractors to solicit and maintain data on the gender, race, and ethnicity of each job applicant for agency monitoring and enforcement purposes. The OFCCP uses these data during "desk audit," the initial stage of a compliance evaluation, to review a contractor's applicant and hiring data and assess whether employment practices comply with equal opportunity obligations. Based on this review, the OFCCP determines whether an on-site investigation is warranted. Because most evaluations conclude at the desk audit stage, it serves as the agency's primary method for identifying potential discrimination patterns.[12] Critically, while

[12] See https://www.federalregister.gov/documents/2005/10/07/05-20176/obligation-to-solicit-race-and-gender-data-for-agency-enforcement-purposes.

federal contractors face audit risk comparable to the Bias Audit Law's requirements, they are not required to disclose their applicant demographic data or audit results publicly.

This institutional setting allows us to distinguish between auditing and disclosure effects. For federal contractors, the Bias Audit Law primarily imposes an additional disclosure obligation, as they already face audit requirements under OFCCP oversight. For non-federal contractors, the law introduces both auditing and disclosure requirements. Therefore, if the disclosure requirement plays a role in driving hiring changes, we should observe a reduction in male hiring among firms that have previously been federal contractors. Conversely, if federal contractors show no significant effect while non-federal contractors exhibit significant reductions in male hiring, the pattern will suggest that disclosure alone is insufficient to drive changes and that the independent auditing requirement serves as the primary mechanism driving hiring diversity improvements.

To test this mechanism, we collect data on federal contractors from the FPDS-NG database, publicly available via USAspending.gov, and classify firms into two groups: federal contractors and non-federal contractors. Table 6 reports the results for these subsamples, with Column (1) [2] for [non-]federal contractors. *Treat* × *Post* is significantly negative only in Column (2), indicating a reduction in male hiring after the enforcement of the Bias Audit Law among firms that have never been federal contractors. These findings suggest that the auditing requirement is likely to be the primary mechanism driving the observed effects.[13]

[Insert Table 6 about here]

[13] A caveat is that federal contractors might differ in their propensity to use AI-based recruiting tools. While descriptive evidence suggests that 55.2% of contractors operate in AI-intensive segments, we cannot fully disentangle this channel with our current data.

6.4 Subsample Analyses: Private and Public firms

Previous studies find that 80% of U.S. companies and nearly all Fortune 500 firms use AI-powered hiring software (Wiessner 2024; Dennison 2024). To examine whether our results are driven solely by these firms, we conduct subsample analyses based on firm public status.

We classify firms as public if they have a GVKEY identifier provided by Revelio Labs. Table 7 presents the results. Column (1) reports estimates for private firms, and Column (2) for public firms.[14] Across all columns, the *Treat*×*Post* interaction terms are statistically significant, indicating that both public and private firms reduce male hiring following implementation of Bias Audit laws. These results suggest that the regulatory impact extends beyond large, publicly traded firms and affects broader segments of the labor market.[15]

[Insert Table 7 about here]

6.5 Subsample Analyses: Job Seniority

We further examine whether the impact of the Bias Audit Law varies by job seniority level. Senior positions often require specific skills and extensive experience, which lowers the pool of eligible candidates and can limit the applicability of standardized automated screening tools. These roles may also rely more heavily on networking and executive search processes that fall outside the scope of algorithmic decision tools. In contrast, junior positions typically involve

[14] In Table 7, the number of observations for private firms is smaller than that for public firms due to data limitations in the merging process. We merge datasets from Revelio Labs, the Individual Database and the Job Posting Database, with the Workforce Dynamics Database separately, which provides total employment information. However, as confirmed by a Revelio Labs representative, not all companies in the Individual or Job Posting Databases can be matched to the Workforce Dynamics Database. This issue is especially common among smaller firms, which may have detailed job posting or individual-level data but lack complete employment records in Workforce Dynamics. As a result, when total employment is included as a control, the sample size for private firms (firms without GVKEY) is substantially reduced. If we exclude total employment as a control and do not merge with the Workforce Dynamics Database, the number of observations in the private firm sample becomes roughly four times larger than that in the public firm sample (firms with GVKEY). Importantly, our findings remain robust, suggesting that the main results are not driven by the inclusion of employment controls or the smaller sample.

[15] In untabulated results, we also examine the heterogeneity of the Bias Audit Law's effects by firm size. We find that the reduction in male hiring following the law's enforcement is concentrated among large and medium-sized firms, which are more likely to adopt AI-based hiring practices.

more standardized hiring processes with clearly defined qualifications and much larger volumes of applications, making them more suitable for AI-based screening (Godfrey, 2025; Jaser et al., 2021). In addition, junior roles entail lower adjustment costs when firms respond to fairness mandates (Chang and Kirgios, 2024; Larcker et al., 2025). Therefore, we expect that the decline in male hiring is primarily concentrated in junior positions.

To test this prediction, we conduct subsample analyses based on seniority levels in Revelio Labs. Table 8 presents the results of estimating separate regressions from entry-level positions (seniority = 1) in column (1) to senior executive positions (seniority = 7) in column (7). The coefficients on *Treat×Post* are significantly negative only in columns (1) through (4). These results suggest that the reduction in male hiring is particularly pronounced in lower to mid-tier positions, likely because AI screening tools are more relevant for these positions or diversity adjustments are easier to implement among these roles.

[Insert Table 8 about here]

This analysis also helps alleviate concerns that NYC firms respond more to national trends—such as corporate diversity fellowships, on-campus recruiting expansions, unionization drives, and skills-based hiring initiatives—due to the city's unique culture. Such initiatives disproportionately target junior-level or lower-level hiring through fellowship programs designed for recent graduates and campus recruiting partnerships. However, our findings hold even for management-level employees (levels 3-4), where these initiatives are much less applicable. The persistence of effects at management levels strengthens confidence that the Bias Audit Law, rather than these confounding events, drives our observed results.

6.6 Bias Audit and Job Vacancy Duration

This section examines the potential costs of Bias Audit Laws proxied by the time offices spend filling open positions (H3). The underlying rationale is that bias audits enhance transparency, improve stakeholder monitoring, and heighten potential reputational and litigation risks for firms whose algorithms display discriminatory patterns. In response, affected firms may devote more time to administrative compliance for third-party audits and proactively adjust their recruitment practices in anticipation of public scrutiny. Such adjustments may include expanding candidate pools to increase diversity, adding supplementary screening steps, or conducting more thorough evaluations before the next audit cycle. Therefore, increased hiring diversity may come at the cost of reduced recruitment efficiency. To assess this trade-off, we investigate how the duration of job vacancies changes following the enforcement of the Bias Audit Law.

The results are presented in Table 9. Column (1) reports OLS estimates with *LogDuration*, the natural logarithm of the average number of days to fill a vacancy in a quarter, as the dependent variable. Column (2) shows Poisson estimates with *Duration*, the average number of days to fill a vacancy in a quarter, as the dependent variable. The coefficients on *Treat*×*Post* are significantly positive in both specifications (0.202, t=19.638 for OLS regression; 0.206, t=21.983 for Poisson regression), indicating that New York offices experienced approximately a 22%-23% increase in average job vacancy duration relative to control offices. The effect is also economically significant. Given the baseline average vacancy duration of 40.2 days, the estimated percentage increase translates to approximately 9 additional days per hire.

[Insert Table 9 about here]

We also examine the dynamic effect of the Bias Audit Law on job vacancy duration. As shown in Figure 4, no significant differences in job vacancy duration are visible between the treatment and control groups in the pre-adoption period, supporting the parallel trend assumption. However, post-adoption, we observe a significant increase.

[Insert Figure 4 about here]

Together with the results on hiring diversity, these findings suggest that while Bias Audit Law increases female hiring consistent with a previous imbalance, it also decreases recruitment efficiency by extending job vacancy duration. We caveat, though, that because this analysis uses Revelio Labs data, it primarily applies to professional and white collar professions.

6.7 Bias Audit and Total Hires

Another potential cost is that companies may reduce recruitment for their New York offices and shift hiring to other locations. To investigate this possibility, we focus on the total number of new hires. The results are presented in Table 10. Column (1) shows OLS regression results, where the dependent variable is *LogHiring*, the natural logarithm of the total number of hires in a quarter. Column (2) presents Poisson regression results, with *Hiring*, the raw count of hires in a quarter, as the dependent variable. The coefficients on *Treat*Post* are insignificant in both columns. Therefore, we find no evidence to suggest a shift in recruitment to other offices following the implementation of the law.

[Insert Table 10 about here]

6.8 Bias Audit and Subsequent Firm Performance: Evidence from Financial Statement Auditors

In this section, we further investigate the consequences of hiring outcomes resulting from the Bias Audit Laws by examining the performance of financial statement audit firms. We focus on audit offices for several reasons. First, office-level audit quality data are available from their client financial statements, enabling direct measurement of performance outcomes. Second, unlike other financial service occupations that are heavily concentrated in New York City, audit offices are geographically distributed across the United States, providing sufficient variation in our treatment and control groups.

The analysis proceeds in two steps. First, we re-examine Equation (1) using the audit firm subsample to confirm that the hiring effects documented in our full sample extend to this specific industry. Next, we analyze how audit quality changes following the implementation of the Bias Audit Laws. We use misstatements of a company's fiscal year-end that are subsequently restated as our primary measure of audit quality, consistent with prior literature and the idea that an auditor's role is to provide assurance that the client's financial statements are free of misstatements (DeFond and Zhang 2014; Aobdia 2019; Aobdia et al. 2024).

The results are presented in Table 11. Panel A reports the effects on hiring diversity and job vacancy duration. We present results at both monthly and quarterly frequencies to ensure robustness across different temporal aggregations. Columns (1) and (3) report effects on *MaleProp*, columns (2) and (4) examine *WhiteProp*, columns (5) and (7) analyze *LogDuration* using OLS, and columns (6) and (8) employ Poisson regressions with Duration. Panel B reports audit quality results proxied by restatements of fiscal year ends in column (1). We also consider quarterly restatements in column (2). Quarterly restatements should be less affected by any effect of the Bias Audit Law, because only annual financial statements are audited, whereas quarterly financial statements are only reviewed by the financial statement auditor.

As shown in Table 11, Panel A, the coefficients on *Treat*×*Post* are with one exception significantly negative for male hiring and significantly positive for job vacancy duration, suggesting that our main results, declines in male hiring and increases in job posting duration, hold within the audit firm subsample. In Panel B, the coefficient on *Treat*×*Post* is negative in column (1), consistent with a positive effect of the Bias Audit Law on audit quality. This result is also consistent with Khavis et al. (2025), who find that a greater proportion of female auditors in an audit office is associated with greater audit quality. In contrast, the results in column (2) are insignificant, consistent with a more limited auditor role for quarterly financial statements.

[Insert Table 11 about here]

Overall, our results provide some evidence that discipline introduced by Bias Audits leads to improvements in audit firms' performance. Thus, these results help answer whether companies truly address fairness issues in recruiting or make arbitrary changes to their processes that do not address fairness issues in recruiting. Under the former scenario, we would expect the newly hired employees to be of higher quality on average, thereby resulting in performance improvements. Under the latter, we would expect unchanged or even decreased performance.

6.9 Bias Audits and Hiring Patterns of Non-White Racial Subgroups

To better understand the impact of Bias Audit Laws on different racial groups, we conduct an additional analysis focusing on non-white subgroups. In the Revelio Labs dataset, race categories are defined as mutually exclusive. We re-estimate our main specification by replacing the dependent variable with the proportion of hires for each subgroup: *BlackProp* (Black hires), *HispanicProp* (Hispanic hires), and *AsianProp* (Asian hires). Online appendix Table A presents the results, with Columns (1)–(3) corresponding to each subgroup. The coefficient on *Treat*×*Post* is significantly positive in Column (1), indicating an increase in Black hiring, while it is significantly negative in Column (2), suggesting a decline in Hispanic hiring. These findings imply that while the overall share of non-white hires may remain stable, Bias Audit Laws may lead to a redistribution of hiring opportunities among non-white racial groups.

6.10 Robustness Tests

We conduct several robustness tests that we report in Table 12. First, to ensure our findings are not sensitive to quarterly aggregation, we re-estimate our models using office-month data in Panel A. This higher-frequency specification provides additional temporal variation and addresses concerns that quarterly aggregation might mask important month-to-month dynamics or create artificial patterns in the data. Second, to account for the possibility that New York's strong commitment to diversity may distinguish it from other labor markets, we narrow our control group in Panel B to offices of the same firms located in other cities within New York State. Third, we replace office and Year-Quarter fixed effects with Firm×Year-Quarter fixed effects in Panel C. This specification provides a more stringent test by comparing treated and control offices within the same firm and time period, effectively controlling for any concurrent firm-level policies or shocks that might affect all offices simultaneously.

Fourth, in Panel D, we address potential concerns about violations of the Stable Unit Treatment Value Assumption (SUTVA), which could occur if control offices are indirectly affected by the treatment. To mitigate this concern, we re-estimate the regression using an alternative control group that includes only offices from firms that have never had an office in New York. As shown in Table 12, our main results, declines in male hiring and increases in job posting duration, hold across all of these robustness checks.[16]

[Insert Table 12 about here]

7. Conclusion

This paper provides the first large-sample empirical evidence on the effects of Bias Audit

[16] In untabulated analyses, we conduct several additional robustness tests. First, to ensure greater comparability between treated and control offices, we restrict the control sample to offices in large metropolitan areas with similar economic characteristics, such as Los Angeles, Chicago, Dallas, Houston, and Washington, D.C. Our main results remain robust. Second, we address concerns arising from the extended timeline between the law's introduction (February 2020) and enforcement (July 2023). During this three-year gap, firms may have gradually adjusted their practices or anticipated the requirement. We use a "clean" comparison period by restricting our sample to pre-introduction quarters (Q1-Q4 2019) and post-enforcement quarters, excluding the 2020-2022 interim period. Our results remain consistent, suggesting that transitory adjustments during implementation do not drive our findings. Then, we test for anticipation effects around the official announcement in Q4 2021. We examine two alternative windows: (1) Q1 2019-Q3 2021 versus Q4 2021-Q2 2023, and (2) Q4 2020-Q3 2021 versus Q4 2021-Q4 2022. We find no significant changes in male hiring around the announcement date, indicating no anticipation effects and confirming that our results reflect responses to actual enforcement.

Law on hiring outcomes. Using a Difference-in-Differences design, we find a decrease in the proportion of male hires following implementation of the Bias Audit Law. We also find that the effects are more pronounced in firms with pre-existing gender imbalances or firms with higher AI intensity. Additionally, the reduction in male hires is observed only among firms that have never been federal contractors, rather than among federal contractors, which is consistent with the auditing mechanism (in contrast with the disclosure mechanism embedded in the law) driving the results. The decline is also salient among lower to mid-tier roles and exists in both public and private companies. We also find some evidence that Bias Audits improve financial statement audit client outcomes by reducing restatements. Overall, our study offers timely insights for policymakers evaluating similar regulations and contributes to the growing literature on AI governance and the evolving role of audits in non-financial domains.

References

Adams, R. B., and D. Ferreira. 2009. Women in the boardroom and their impact on governance and performance. *Journal of Financial Economics* 94 (2): 291–309.

Adhikari, B. K., A. Agrawal, and J. Malm. 2019. Do women managers keep firms out of trouble? Evidence from corporate litigation and policies. *Journal of Accounting and Economics* 67 (1): 202–225.

Ahern, K. R., and A. K. Dittmar. 2012. The Changing of the Boards: The Impact on Firm Valuation of Mandated Female Board Representation *. *The Quarterly Journal of Economics* 127 (1): 137–197

Ahn, J., R. Hoitash, U. Hoitash, and E. Krause. 2025. Diversity and Career Trajectories: Evidence from LinkedIn Data on Race, Ethnicity, and Gender in Auditing. *The Accounting Review* 100 (4): 1–31.

Aobdia, D. 2019. Do practitioner assessments agree with academic proxies for audit quality? Evidence from PCAOB and internal inspections. *Journal of Accounting and Economics* 67(1): 144-174.

Aobdia, D., P. Choudhary, and N. Newberger. 2024. The economics of audit production: what matters for audit quality? An empirical analysis of the role of midlevel managers within the audit firm. *The Accounting Review* 99(2): 1-29.

Aobdia, D., G. Köchling, P. Limbach, and A. Yoon. 2025. Emissions Restatements After the SEC Climate Proposal: Evidence from Carbon Disclosure Project Filings. Working paper.

Arrow, K., Ashenfelter, O. and Rees, A., 1973. Discrimination in labor markets. *The Theory of Discrimination*, pp.3-33.

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

BABL AI. 2022. Public Comment on Proposed Rules Related to Local Law 144 (Automated Employment Decision Tools). New York City Department of Consumer and Worker Protection, June 3. Available at: https://www.nyc.gov/assets/dca/downloads/pdf/about/PublicComments-Proposed-Rules-Related-to-LocalLaw202of2019-LocalLaw1144of2021-LocalLaw37of2022.pdf.

Baker, A. C., D. F. Larcker, C. G. McClure, D. Saraph, and E. M. Watts. 2024. Diversity Washing. *Journal of Accounting Research* 62 (5): 1661–1709.

Barocas, S., and A. D. Selbst. 2016. Big Data's Disparate Impact. *California Law Review* 104: 671.

Becker, G. S. 1957. *The Economics of Discrimination*. University of Chicago Press.

Bian, B., J. Li, and K. Li. 2025. Does Mandating Women on Corporate Boards Backfire? SSRN Scholarly Paper. Rochester, NY: Social Science Research Network.

Blankespoor, E., E. deHaan, and Q. Li. 2025. Generative AI in Financial Reporting. SSRN Scholarly Paper. Rochester, NY: Social Science Research Network.

Billings, M. B., A. Klein, and Y. C. Shi. 2022. Investors' response to the #MeToo movement: does corporate culture matter? *Review of Accounting Studies* 27 (3): 897–937.

Bonfiglioli, A., R. Crinò, M. Filomena, and G. Gancia. 2025. Comparative Advantage in AI-Intensive Industries: Evidence from US Imports. SSRN Scholarly Paper. Rochester, NY: Social Science Research Network.

Bourveau, T., J. Brendel, and J. Schoenfeld. 2024. Decentralized Finance (DeFi) assurance: early evidence. *Review of Accounting Studies* 29 (3): 2209–2253.

Bourveau, T., X. Gao, and O.-K. Hope. 2025. The Impact of Disclosure on Diversity: Evidence from the Canada Business Corporations Act. SSRN Scholarly Paper. Rochester, NY: Social Science Research Network.

Bradshaw, M. T., C. Ma, B. Yost, and Y. Zou. 2025. Generative AI Use by Capital Market Information Intermediaries: Evidence from Seeking Alpha. SSRN Scholarly Paper. Rochester, NY: Social Science Research Network.

Breuer, M., W. Cai, A. Le, and F. Vetter. 2024. Minority Representation at Work. SSRN Scholarly Paper. Rochester, NY: Social Science Research Network.

Cai, W., Y. Chen, S. Rajgopal, and L. Azinovic-Yang. 2024. Diversity targets. *Review of Accounting Studies* 29 (3): 2157–2208.

Celiktutan, B., R. Cadario, and C. K. Morewedge. 2024. People see more of their biases in algorithms. *Proceedings of the National Academy of Sciences* 121 (16): e2317602121.

Chang, A., X. Dong, X. Martin, and C. Zhou. 2025. AI (ChatGPT) Democratization and Trading Inequality. SSRN Scholarly Paper. Rochester, NY: Social Science Research Network.

Chen, C.-W., and L. Y. Li. 2023. Is hiring fast a good sign? The informativeness of job vacancy duration for future firm profitability. *Review of Accounting Studies* 28 (3): 1316–1353.

Chang, E. H., and E. L. Kirgios. 2024. Demographic "Stickiness": The Demographic Identity of Departing Group Members Influences Who Is Chosen to Replace Them. *Management Science* 70 (7): 4236–4259.

Chen, Z. 2023. Ethics and discrimination in artificial intelligence-enabled recruitment practices. *Humanities and Social Sciences Communications* 10 (1): 567.

Choi, J. H., J. Pacelli, K. M. Rennekamp, and S. Tomar. 2023. Do jobseekers value diversity information? Evidence from a field experiment and human capital disclosures. *Journal of Accounting Research* 61 (3): 695–735.

Choi, J. H., and C. Xie. 2025. Human + AI in Accounting: Early Evidence from the Field. SSRN Scholarly Paper. Rochester, NY: Social Science Research Network.

Dastin, J. 2018. Insight- Amazon scraps secret AI recruiting tool that showed bias against women. Reuters, October 10.

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

DeFond, M., and J. Zhang. 2014. A review of archival auditing research. *Journal of Accounting and Economics* 58 (2–3): 275–326.

Dennison, K. 2024. Could lawsuits against AI lead to a shift in job searching? Forbes, March 21. https://www.forbes.com/sites/karadennison/2024/03/21/could-lawsuits-against-ai-lead-to-a-shift-in-job-searching/.

Dong, T. 2022. Gender Salary Gap in the Auditing Profession: Trend and Explanations. *European Accounting Review* 33(2): 617–645.

Francis, J. R. 2004. What do we know about audit quality? *The British Accounting Review* 36 (4): 345–368.

Gassam Asare, J. 2025. What the Workday lawsuit reveals about AI bias – and how to prevent it. Forbes, June 23.

Gipper, B., S. Ross, and S. X. Shi. 2025. ESG Assurance in the United States. SSRN Scholarly

Paper. Rochester, NY: Social Science Research Network.
Glum, J. 2023. Summer of strikes: why so many workers are walking off the job. Money, July 21. https://money.com/worker-strikes-summer-2023/
Godfrey, A. 2025. Understanding public perceptions towards automated decision-making in recruitment. *Revealing Reality*.
Goodman, R. 2018. Why Amazon's automated hiring tool discriminated against women. ACLA News & Commentary, October 12.
Hickock, M. 2022. NYC Bias Audit Law: Clock ticking for Employers and HR Talent Technology Vendors. Credo AI. August 15.
Hilliard, A., A. Gulley, A. Koshiyama, and E. Kazim. 2024. Bias audit laws: how effective are they at preventing bias in automated employment decision tools? International Review of Law, Computers & Technology, 1-17.
Jaser, Z., D. Petrakaki, R. Starr, and E. Oyarbide-Magaña. 2021. Automated job interviews and the implications for young jobseekers.
Khavis, J., A. Sheneman, and B. Szerwo. 2025. Does gender composition of audit teams matter? An examination of audit quality and audit cost. *Review of Accounting Studies*, forthcoming.
Kim, P. T. 2016. Data-Driven Discrimination at Work. *William & Mary Law Review* 58: 857.
Kim, P. T. 2017. Auditing Algorithms for Discrimination. *University of Pennsylvania Law Review Online* 166: 189.
Knechel, W.R., S. Maex, and H.J. Park. 2025. The role of auditor reputation in an emerging audit marketplace: Evidence from Decentralized Finance (DeFi). *Management Science*, forthcoming.
Kroll, J., J. Huey, S. Barocas, E. Felten, J. Reidenberg, D. Robinson, and H. Yu. 2017. Accountable Algorithms. *University of Pennsylvania Law Review* 165 (3): 633.
Lam, K., B. Lange, B. Blili-Hamelin, J. Davidovic, S. Brown, and A. Hasan. 2024. A Framework for Assurance Audits of Algorithmic Systems. In *The 2024 ACM Conference on Fairness, Accountability, and Transparency*, 1078–1092.
Larcker, D. F., C. McClure, S. X. Shi, and E. M. Watts. 2025. The Limited Corporate Response to DEI Controversies. SSRN Scholarly Paper. Rochester, NY: Social Science Research Network.
Lawrence, A., M. Minutti-Meza, and P. Zhang. 2011. Can Big 4 versus Non-Big 4 Differences in Audit-Quality Proxies Be Attributed to Client Characteristics? *The Accounting Review* 86 (1): 259–286.
Leicht-Deobald, U., T. Busch, C. Schank, A. Weibel, S. Schafheitle, I. Wildhaber, and G. Kasper. 2019. The Challenges of Algorithm-Based HR Decision-Making for Personal Integrity. *Journal of Business Ethics* 160 (2): 377–392.
Levy, B. 2025. Caution Ahead: Numerical Reasoning and Look-ahead Bias in AI Models. SSRN Scholarly Paper. Rochester, NY: Social Science Research Network.
Lins, K. V., L. Roth, H. Servaes, and A. Tamayo. 2024. Sexism, culture, and firm value: evidence from the Harvey Weinstein scandal and the# MeToo movement. *Journal of Accounting Research* 62 (5): 1989–2035.
Liu, T., C. A. Makridis, P. Ouimet, and E. Simintzi. 2023. The Distribution of Nonwage Benefits: Maternity Benefits and Gender Diversity. *The Review of Financial Studies* 36 (1): 194–234
Lockhart, J. W., M. M. King, and C. Munsch. 2023. Name-based demographic inference and the

unequal distribution of misrecognition. *Nature Human Behaviour* 7 (7): 1084–1095.

Matsa, D. A., and A. R. Miller. 2013. A Female Style in Corporate Leadership? Evidence from Quotas. *American Economic Journal: Applied Economics* 5 (3): 136–169.

McColl, R., and M. Michelotti. 2019. Sorry, could you repeat the question? Exploring video-interview recruitment practice in HRM. *Human Resource Management Journal* 29 (4): 637–656.

Milkman, R., and J. van der Naald. 2023. The State of the Unions 2023: A Profile of Organized Labor in New York City, New York State, and the United States. *Publications and Research*.

Moschella, D. 2022. *AI Bias Is Correctable. Human Bias? Not So Much*.

Munoko, I., H. L. Brown-Liburd, and M. Vasarhelyi. 2020. The Ethical Implications of Using Artificial Intelligence in Auditing. *Journal of Business Ethics* 167 (2): 209–234.

Oliveira, M., and S. Zhang. 2022. The trends and determinants of board gender and age diversities. *Finance Research Letters* 46: 102798.

Peng, A., B. Nushi, E. Kiciman, K. Inkpen, and E. Kamar. 2022. Investigations of Performance and Bias in Human-AI Teamwork in Hiring. *Proceedings of the AAAI Conference on Artificial Intelligence* 36 (11): 12089–12097.

Prince, A. E. R., and D. Schwarcz. 2019. Proxy Discrimination in the Age of Artificial Intelligence and Big Data. *Iowa Law Review* 105: 1257.

Raghavan, M., S. Barocas, J. Kleinberg, and K. Levy. 2020. Mitigating bias in algorithmic hiring: evaluating claims and practices. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 469–481. FAT* '20. New York, NY, USA: Association for Computing Machinery.

Raji, I. D., and J. Buolamwini. 2019. Actionable Auditing: Investigating the Impact of Publicly Naming Biased Performance Results of Commercial AI Products. In *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, 429–435. AIES '19. New York, NY, USA: Association for Computing Machinery.

Suresh, H., and J. V. Guttag. 2021. A Framework for Understanding Sources of Harm throughout the Machine Learning Life Cycle. In *Equity and Access in Algorithms, Mechanisms, and Optimization*, 1–9.

Sutherland, A. G., M. Uckert, and F. W. Vetter. 2024. Occupational licensing and minority participation in professional labor markets. *Journal of Accounting Research* 62 (2): 453–503.

Van den Broek, E., A. Sergeeva, and M. Huysman. 2021. When the Machine Meets the Expert: An Ethnography of Developing AI for Hiring. *MIS Quarterly* 45 (3).

Wade, H. 2022. NYC Rules. https://rules.cityofnewyork.us/rule/automated-employment-decision-tools-2/.

Watts, R. L., and J. L. Zimmerman. 1983. Agency Problems, Auditing, and the Theory of the Firm: Some Evidence. *The Journal of Law and Economics* 26 (3): 613–633.

Wiessner. 2023. Tutoring firm settles US agency's first bias lawsuit involving AI software. https://www.reuters.com/legal/tutoring-firm-settles-us-agencys-first-bias-lawsuit-involving-ai-software-2023-08-10/.

Wiessner, D. 2024. Workday accused of facilitating widespread bias in novel AI lawsuit. *Reuters*, February 21, sec. Transactional.

Wooldridge, J.M. 2001. Econometric analysis of cross section and panel data. The MIT Press.

Appendix Table A: Variable Definition

Variable Name	Variable Description
MaleProp	The proportion of male employees hired in the office during the quarter
WhiteProp	The proportion of white employees hired in the office during the quarter
WhiteMaleProp	The proportion of white male employees hired in the office during the quarter
NonWhiteMaleProp	The proportion of non-white male employees hired in the office during the quarter
PropAgeUnder25	The proportion of employees under 25 years old hired in the office during the quarter
PropAge25To34	The proportion of employees aged 25-34 hired in the office during the quarter
PropAge35To44	The proportion of employees aged 35-44 hired in the office during the quarter
PropAge45To54	The proportion of employees aged 45-54 hired in the office during the quarter
PropAge55Plus	The proportion of employees over 55 years old hired in the office during the quarter
LogDuration	The natural logarithm of the average number of days to fill a job vacancy in the office during the quarter
LogHiring	The natural logarithm of the total number of employees hired in the office during the quarter
LogEmpNum	The natural logarithm of the number of employees in the office during the quarter
HighMaleFirm	Indicator variable equal to 1 if the firm's proportion of male hires before the regulation was above the median, and 0 otherwise.
HighMaleInd	Indicator variable equal to 1 if a firm is in a male-dominated industry, defined as industries where the female employment share is below the overall average across all industries, based on the Labor Force Statistics from the Current Population Survey (2024)
HighAIFirm	Indicator variable equal to one if the office is among the top AI-investing firms, proxied by whether the office has hired AI-skilled workers before the Bias Audit Law, and zero otherwise, following Babina et al. (2024).
HighAIInd	Indicator variable equal to 1 if the company's industry is one of the highest AI-investment sectors, based on Table 2 in Bonfiglioli et al. (2025), and 0 otherwise.
BlackProp	The proportion of Black employees hired in the office during the quarter
HispanicProp	The proportion of Hispanic employees hired in the office during the quarter
AsianProp	The proportion of Asian employees hired in the office during the quarter
RestateQtr	Indicator variable equal to 1 if the company's quarterly financial statements are subsequently restated, and 0 otherwise
RestateAnn	Indicator variable equal to 1 if the company's annual financial statements are subsequently restated, and 0 otherwise.
Treat	Indicator variable equal to 1 if the office is located in New York City, and 0 otherwise.
Post	Indicator variable equal to 1 for quarters on or after 2023 Q3 when the bias-audit law became effective, and 0 for quarters before
Size	The natural logarithm of market value of equity. The sample is at the audit firm-client-quarter level.

Variable Name	Variable Description
MTB	Market to book ratio. The number of common shares outstanding multiplied by stock price divided by the book value of equity. The sample is at the audit firm-client-quarter level.
Lev	Total liabilities divided by total assets. The sample is at the audit firm-client-quarter level.
ROA	Net income before extraordinary items divided by average total assets. The sample is at the audit firm-client-quarter level.
Loss	A dummy variable that equals one if net income is negative and zero otherwise. The sample is at the audit firm-client-quarter level.
Big4	A dummy variable that takes value one if the firm's current auditor is a big four auditor, and zero otherwise. The sample is at the audit firm-client-quarter level.
Ln_auditfee	The natural logarithm of audit fees.
CFO	Cash flow from operations deflated by beginning assets.
Auditor_firstyear	Indicator variable equal to one when the client audited by the audit for the first time in year t, zero otherwise.
M&A	Indicator variable equal to one when the client has merger or acquisition activity, zero otherwise.
Num	Number of business segments
Foreign	Indicator variable equal to one when the client has foreign operations, zero otherwise.
Disc_ops	Indicator variable equal to one when the client has discontinued operations zero otherwise.

Appendix B: A Sample of Audit Report

Appendix B presents a sample audit report issued by BABL for the bias audit results of Eightfold.

Bias Audit for New York City Local Law 144
Prepared by BABL AI Inc. | 03/21/2025
Letter from the Lead Auditor | Summary | Conclusions | Findings

babl

Letter from the Lead Auditor

From: **Shea Brown**
Lead Auditor
BABL AI Inc.
sheabrown@bablai.com

To: **Eightfold AI Inc.**
2625 Augustine Drive
Suite 601
Santa Clara, CA 95054

Re: **Audit Opinion on Eightfold AI Inc.'s Eightfold Matching Model**

03/21/2025

We have independently audited the bias testing assertions and related documentary evidence of Eightfold AI Inc. (the "Company") as of 03/21/2025, presented to BABL AI in relation to Company's Eightfold Matching Model in accordance with the criteria and audit methodology set forth in this report. The goals of this audit are to:

1. Determine whether the bias testing methodologies, controls, and procedures performed by Company satisfy the audit criteria (see Findings)
2. Obtain reasonable assurance as to whether the statements made by the Company, including the summary of bias testing results presented in this report, are free from material misstatement, whether due to fraud or error.

Note that the criteria presented in this report were constructed specifically to address the requirements of a "bias audit" outlined in NYC Local Law No. 144 of 2021. The model was audited as though it were an automated employment decision tool (AEDT) under NYC Local Law No. 144 of 2021, but we do not make any determination whether the model is, in fact, an AEDT under this law.

Company Responsibilities

It is the responsibility of Company representatives to ensure that bias testing and related procedures comply with the criteria outlined in this report. The Company representatives are responsible for ensuring that the documents submitted are fairly presented and free of misrepresentations, providing all resources and personnel needed to ensure an effective and efficient audit process, and providing access to evidential material as requested by the auditors.

BABL AI Responsibilities

It is the responsibility of the lead auditor to express an opinion on the Company's assertions related to the bias testing of the model. In light of the current absence of generally accepted standards for the auditing of algorithms and autonomous systems, our examination was conducted in accordance with the standards and normative references outlined in this report.

Those standards require that we plan and perform audit procedures to obtain reasonable assurance about whether the assertions referred to above 1) satisfy the audit criteria and 2) are free of material misstatement, whether due to error or fraud. Within the scope of our engagement, we performed amongst others the following procedures:

- Inspection of submitted documents and external documentation
- Interviewing Company employees to gain an understanding of the process for determining the disparate impact and risk assessment results
- Observation of selected analytical procedures used in Company's bias testing
- Inspection of the select samples of the bias testing data
- Inquiry of personnel responsible for governance and oversight of the bias testing and risk assessment

We believe that the procedures performed provide a reasonable basis for our opinion.

Independence

Our role as an independent auditor conforms to ForHumanity and Sarbanes-Oxley definitions of Independence. Fees associated with this contract are for the provision of the service to assess compliance. The payment of fees is unrelated to the decision rendered. Our decision is grounded solely in the criteria presented below.

Opinion

In our opinion, based on the procedures performed and the evidence received to obtain assurance, the bias testing and results presented by Company, as of 03/21/2025, is prepared, in all material respects, in accordance with the criteria outlined below.

Sincerely,

Shea Brown

Shea Brown
Lead Auditor, BABL AI Inc.

Non-intersectional, Gender, sorted by Scoring rate

	N applicants	Scoring rate	Impact ratio
Female	116	0.509	1.000
Male	739	0.491	0.966

Non-intersectional, Race/ethnicity, sorted by Scoring rate

	N applicants	Scoring rate	Impact ratio[7]
Black or African American	44	0.545	1.000
Hispanic or Latino	48	0.521	0.955
White	174	0.506	0.927
Asian	589	0.484	0.887

Intersectionals

			N applicants	Scoring rate	Impact ratio[8]
Hispanic or Latino	Male		43	0.512	0.853
	Female		5	0.600	1.000
Non-Hispanic or Latino	Male	White	151	0.503	0.839
		Asian	506	0.482	0.804
		Black or African American	39	0.538	0.897
		Native American or Alaskan Native	0	0.000	N/A
		Native Hawaiian or Pacific Islander	0	0.000	N/A
		Two or more races	0	0.000	N/A
	Female	Asian	83	0.494	0.823
		White	23	0.522	0.870
		Black or African American	5	0.600	1.000
		Native American or Alaskan Native	0	0.000	N/A

Figure 1: Timeline of the regulation

Figure 1 presents the timeline of the New York City Bias Audit Law. The figure is sourced from https://www.nycbiasaudit.com/.

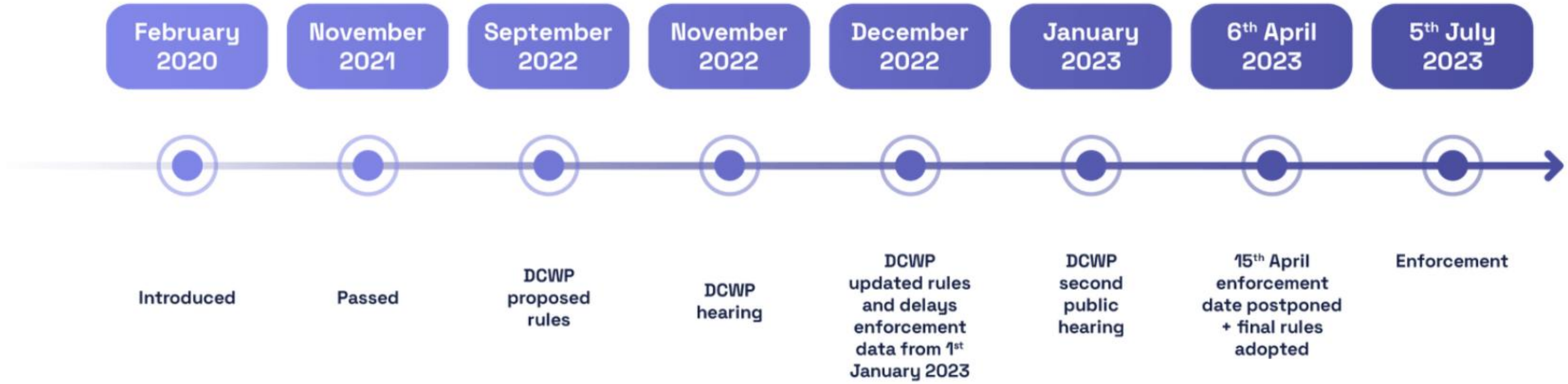
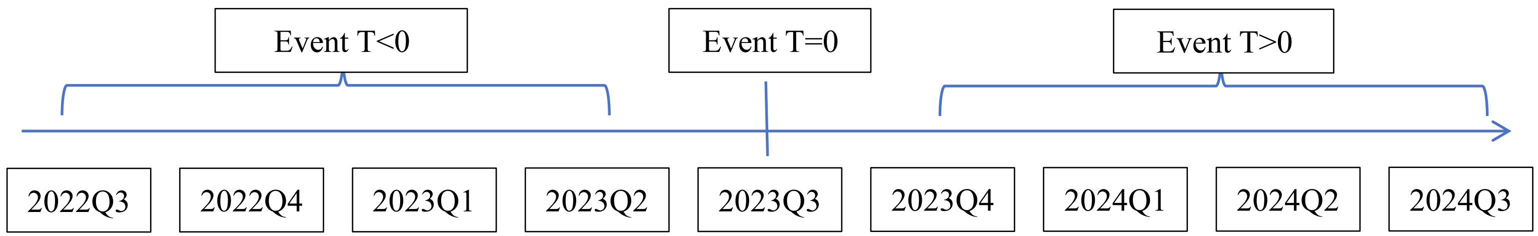


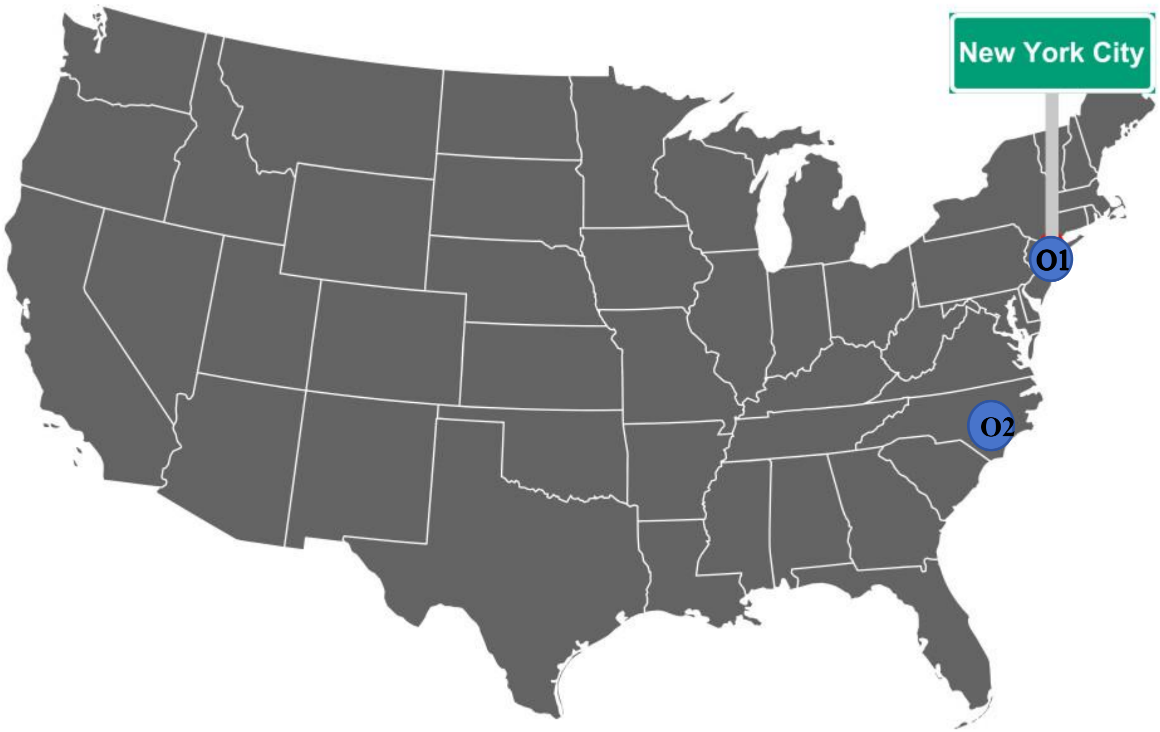
Figure 2: Research Design

Figure 2 presents our research design. Panel A presents our sample window, starting from 2022 Q3 and ending with 2024 Q3. We use 2023 Q3, when the Bias Audit Law was enforced, as the event quarter (Event T= 0). Panel B presents a graphical summary of the identification strategy. Suppose Office 1 is located in New York City, while Office 2, from the same company, is located outside of New York City. Then Office 1 is in the treatment group while Office 2 is in the control group in our empirical design.

Panel A: Sample window



Panel B: Treatment and control group



O1 (Treatment group): the office located in New York

O2 (Control group): the office of the same company located outside New York

Figure 3: Bias Audit Law and Male Proportion of Employment Hiring in Event-Time

Figure 3 displays the dynamic effect of the Bias Audit Law on the male hiring proportion. The unit of observation is at the office-quarter level. The dependent variable is *MaleProp*, the proportion of male employees hired. The model replaces the "Post" indicator with a set of event-time dummies, each measured relative to the benchmark quarter t – 1 (2023 Q2). Solid points plot the estimated coefficients, and vertical lines indicate 90 percent confidence intervals.

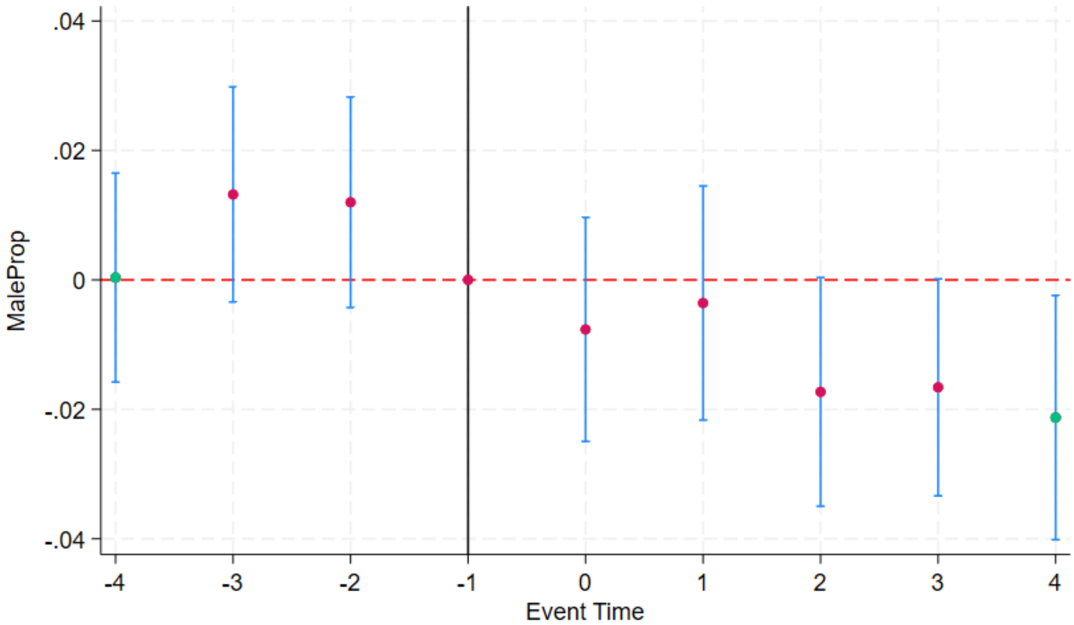


Figure 4: Bias Audit Law and Job Vacancy Duration in Event-Time

Figure 4 displays the dynamic effect of the Bias Audit Law on the job vacancy duration. The unit of observation is at the office-quarter level. The dependent variable is *LogDuration*, the natural logarithm of the average number of days to fill a job vacancy in the office during the quarter. The model replaces the "Post" indicator with a set of event-time dummies, each measured relative to the benchmark quarter t – 1 (2023 Q2). Solid points plot the estimated coefficients, and vertical lines indicate 90 percent confidence intervals.

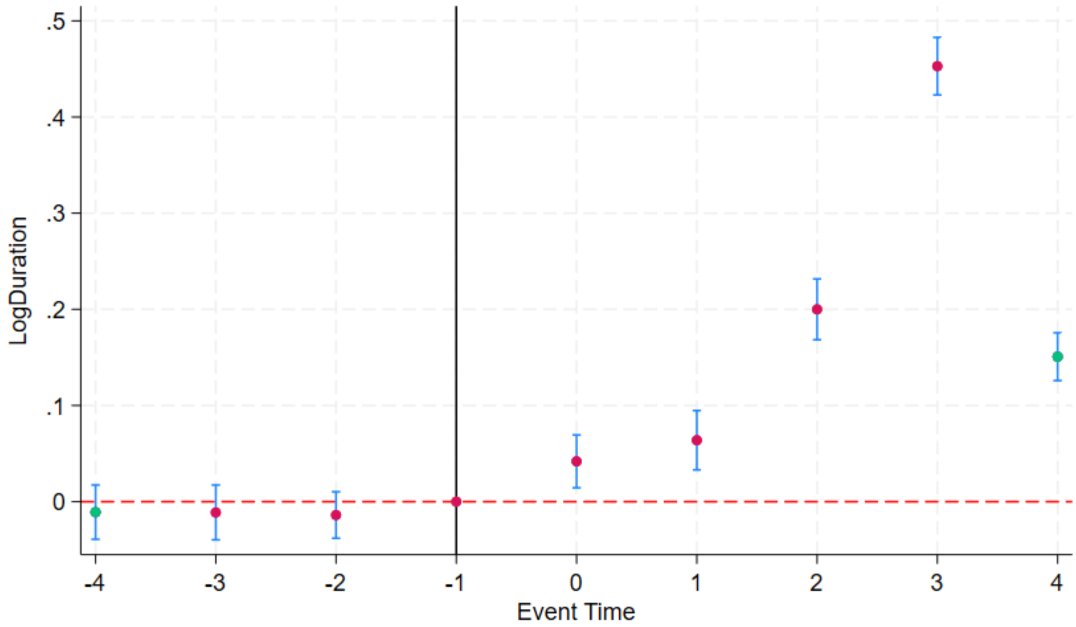


Table 1: Summary statistics

Table 1 presents the summary statistics for our main sample. All variables are defined in Appendix Table A.

	N	Mean	Median	SD	P25	P75
MaleProp	556,511	0.571	0.609	0.398	0.125	1.000
WhiteProp	556,511	0.701	0.889	0.365	0.500	1.000
WhiteMaleProp	556,511	0.403	0.333	0.393	0.000	0.750
NonWhiteMaleProp	556,511	0.167	0.000	0.293	0.000	0.250
PropAgeUnder25	416,487	0.349	0.200	0.397	0.000	0.667
PropAge25To34	416,487	0.347	0.250	0.384	0.000	0.571
PropAge35To44	416,487	0.180	0.000	0.314	0.000	0.250
PropAge45To54	416,487	0.087	0.000	0.232	0.000	0.000
PropAge55Plus	416,487	0.037	0.000	0.156	0.000	0.000
LogDuration	716,495	3.551	3.597	0.566	1.386	4.812
Duration	716,495	40.278	37.000	21.415	27.000	49.000
LogHiring	556,511	1.308	1.099	0.802	0.693	1.609
hiring	556,511	5.569	2.000	22.019	1.000	4.000
LogEmpNum	556,511	4.757	4.682	1.806	3.477	6.003
BlackProp	556,511	0.097	0.000	0.233	0.000	0.000
HispanicProp	556,511	0.103	0.000	0.240	0.000	0.024
AsianProp	556,511	0.094	0.000	0.232	0.000	0.000

Table 2: Effect of the Bias Audit Law on Hiring Diversity

Table 2 examines the effect of the Bias Audit Law on hiring diversity, using office-quarter observations. In Panel A, columns (1)–(2) report results for the proportion of male hires (*MaleProp*), and columns (3)–(4) for the proportion of white hires (*WhiteProp*). Panel B shows the proportion of white male hires (*WhiteMaleProp*) in column (1) and non-white male hires (*NonWhiteMaleProp*) in column (2). The primary explanatory variable is the interaction term between *Treat* and *Post*, where *Treat* equals one for offices located in New York City and *Post* equals one for the quarter the law takes effect and thereafter. Standard errors are clustered at the office level. T-statistics are reported below the coefficients. ***, **, and * correspond to 1%, 5% and 10% significance level. Detailed definitions of all variables are provided in Table A.

Panel A: Proportion of Male Hires and Proportion of White Hires

	(1)	(2)	(3)	(4)
DV =	*MaleProp*	*MaleProp*	*WhiteProp*	*WhiteProp*
Treat* × *Post	**-0.019*** **	**-0.019*** **	0.001	0.000
	(-3.640)	**(-3.660)**	(0.110)	(0.098)
LogEmpNum		-0.005***		-0.003**
		(-3.352)		(-2.167)
Office FE	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes
R^2	0.374	0.374	0.380	0.380
N. of Obs.	556,511	556,511	556,511	556,511

Panel B: Proportion of White Male Hires and Proportion of Non-White Male Hires

	(1)	(2)
DV =	*WhiteMaleProp*	*NonWhiteMaleProp*
Treat* × *Post	**-0.009***	**-0.010*** **
	(-1.698)	**(-2.405)**
LogEmpNum	-0.004***	-0.000
	(-3.121)	(-0.287)
Office FE	Yes	Yes
Year-Quarter FE	Yes	Yes
R^2	0.355	0.342
N. of Obs.	556,511	556,511

Table 3: Placebo Test – Effect of the Bias Audit Law on Age Group Hiring

Table 3 presents the results of a placebo test examining the effect of the Bias Audit Law on age‑group hiring shares, using office-quarter observations. Columns (1)-(5) use, respectively, the proportion of hires aged under 25 (*PropAgeUnder25*), 25–34 (*PropAge25To34*), 35–44 (*PropAge35To44*), 45–54 (*PropAge45To54*), and 55 and over (*PropAge55Plus*) as the dependent variables. The primary explanatory variable is the interaction term between *Treat* and *Post*, where *Treat* equals one for offices located in New York City and *Post* equals one for the quarter the law takes effect and thereafter. Standard errors are clustered at the office level. T-statistics are reported below the coefficients. ***, **, and * correspond to 1%, 5% and 10% significance level. Detailed definitions of all variables are provided in Table A.

	(1)	(2)	(3)	(4)	(5)
DV =	*PropAgeUnder25*	*PropAge25To34*	*PropAge35To44*	*PropAge45To54*	*PropAge55Plus*
Treat* × *Post	-0.003	0.006	0.001	-0.005	0.001
	(-0.681)	(1.028)	(0.162)	(-1.393)	(0.607)
LogEmpNum	0.016***	-0.001	-0.007***	-0.006***	-0.002**
	(9.927)	(-0.621)	(-4.997)	(-5.953)	(-2.511)
Office FE	Yes	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes
R^2	0.446	0.301	0.296	0.292	0.280
N. of Obs.	416,487	416,487	416,487	416,487	416,487

Table 4: Cross-Sectional Tests: Heterogeneous Effects by Pre-Law Male Hiring Share

Table 4 examines whether offices and firms with above-median male hiring shares before the Bias Audit Law experienced stronger effects on the male proportion of hires following the law's enactment, using office-quarter observations. Columns (1)–(2) use the male hiring share (*MaleProp*) as the dependent variable. In Columns (1), the primary explanatory variable is the triple interaction *Treat* × *Post* × *HighMaleFirm*, where *HighMaleFirm* equals one if a firm's pre-law male hire share is above the sample median. In Columns (2), the triple interaction is *Treat* × *Post* × *HighMaleInd*, where *HighMaleInd* equals one if a firm is in a male-dominated industry, defined as industries where the female employment share is below the overall average across all industries. *Treat* equals one for offices located in New York City, and *Post* equals one for the quarter the law takes effect and thereafter. Standard errors are clustered at the office level. T-statistics are reported below the coefficients. ***, **, and * correspond to 1%, 5% and 10% significance level. Detailed definitions of all variables are provided in Appendix Table A.

	(1)	(2)
DV =	*MaleProp*	
Treat* × *Post* × *HighMaleFirm	**-0.027****	
	(-2.351)	
Treat × *Post*	-0.001	-0.009
	(-0.120)	(-1.316)
Post × *HighMaleFirm*	-0.012***	
	(-5.579)	
LogEmpNum	-0.004***	-0.005***
	(-3.170)	(-3.308)
Treat* × *Post* × *HighMaleInd		**-0.017***
		(-1.655)
Post × *HighMaleInd*		-0.004**
		(-2.151)
Office FE	Yes	Yes
Year-Quarter FE	Yes	Yes
R^2	0.374	0.374
N. of Obs.	554,937	555,303

Table 5. Cross-sectional Tests: Heterogeneous Effects in AI-Intensive Firms/Industries

Table 5 examines whether firms with higher AI intensity experienced stronger effects on the proportion of male hires following the Bias Audit Law, using office-quarter observations. Columns (1) and (2) use the male hiring share (*MaleProp*) as the dependent variable. In columns (1), the primary explanatory variable is the triple interaction *Treat* × *Post* × *HighAIFirm*, where *HighAIFirm* equals one if the office is among the top AI-investing firms, proxied by whether the office has hired AI-skilled workers before the Bias Audit Law, and zero otherwise, following Babina et al. (2024). In columns (2), the primary explanatory variable is the triple interaction *Treat* × *Post* × *HighAIInd*, where *HighAIInd* equals one if the company's industry is one of the top AI-investing sectors, based on Table 2 in Bonfiglioli et al. (2025). *Treat* equals one for offices located in New York City, and *Post* equals one for the quarter the law takes effect and thereafter. Standard errors are clustered at the office level. T-statistics are reported below the coefficients. ***, **, and * correspond to 1%, 5% and 10% significance level. Detailed definitions of all variables are provided in Appendix Table A.

	(1)	(2)
DV =	*MaleProp*	
Treat* × *Post* × *HighAIFirm	**-0.027***	
	(-1.935)	
Treat × *Post*	-0.002	-0.017***
	(-0.187)	(-3.159)
Post × *HighAIFirm*	0.001	
	(0.375)	
LogEmpNum	-0.005***	-0.004***
	(-2.940)	(-3.177)
Treat* × *Post* × *HighAIInd		**-0.069****
		(-2.472)
Post × *HighAIInd*		-0.010
		(-1.528)
Office FE	Yes	Yes
Year-Quarter FE	Yes	Yes
R^2	0.366	0.374
N. of Obs.	342,668	555,303

Table 6. Subsample Analyses: Federal contractor and Nonfederal contractor

Table 6 examines the effect of the Bias Audit Law on the proportion of male hires for the federal contractor and nonfederal contractor firm subsamples, using office-quarter observations. Columns (1) and (2) use *MaleProp*, the share of male hires, as the dependent variable. Columns (1) reports estimates for federal contractor firms, and Columns (2) for nonfederal contractor firms. The primary explanatory variable is the interaction term between *Treat* and *Post*, where *Treat* equals one for offices located in New York City and *Post* equals one for the quarter the law takes effect and thereafter. Standard errors are clustered at the office level. T-statistics are reported below the coefficients. ***, **, and * correspond to 1%, 5% and 10% significance level. Detailed definitions of all variables are provided in Appendix Table A.

	(1) Federal contractor	(2) Nonfederal contractor
DV =	*MaleProp*	
Treat* × *Post	**-0.007**	**-0.021*****
	(-0.521)	**(-3.670)**
LogEmpNum	-0.008**	-0.004**
	(-2.422)	(-2.536)
Office FE	Yes	Yes
Year-Quarter FE	Yes	Yes
R^2	0.355	0.377
N. of Obs.	113,521	442,990

Table 7. Subsample Analyses: Private and Public Firms

Table 7 examines the effect of the Bias Audit Law on the proportion of male hires for the private and public firm subsamples, using office-quarter observations. Columns (1) and (2) use *MaleProp*, the share of male hires, as the dependent variable. Columns (1) reports estimates for private firms (firms without GVKEY), and Columns (2) for public firms (firms with GVKEY). The primary explanatory variable is the interaction term between *Treat* and *Post*, where *Treat* equals one for offices located in New York City and *Post* equals one for the quarter the law takes effect and thereafter. Standard errors are clustered at the office level. T-statistics are reported below the coefficients. ***, **, and * correspond to 1%, 5% and 10% significance level. Detailed definitions of all variables are provided in Appendix Table A.

	(1) Private Firms	(2) Public Firms
DV =	*MaleProp*	
Treat* × *Post	**-0.020****	**-0.019*****
	(-2.182)	**(-2.934)**
LogEmpNum	-0.003	-0.005***
	(-1.263)	(-2.901)
Office FE	Yes	Yes
Year-Quarter FE	Yes	Yes
R^2	0.387	0.369
N. of Obs.	156,371	400,140

Table 8: Subsample Analyses: Job Seniority

Table 8 examines how the Bias Audit Law affected male hiring shares and vacancy duration by job seniority. The dependent variable *MaleProp* indicates the proportion of male employees hired. The unit of observation is at the office-quarter level. This table reports results for each Revelio Labs seniority tier (1 = entry, 2 = junior, 3 = associate, 4 = manager, 5 = director, 6 = executive, 7 = senior executive). *Treat* equals one for offices located in New York City, and *Post* equals one for the quarter the law takes effect and thereafter. Standard errors are clustered at the office level. T-statistics are reported below the coefficients. ***, **, and * correspond to 1%, 5% and 10% significance level. Detailed definitions of all variables are provided in Table A.

	(1) Seniority = 1	(2) Seniority = 2	(3) Seniority = 3	(4) Seniority = 4	(5) Seniority = 5	(6) Seniority = 6	(7) Seniority = 7
DV =				*MaleProp*			
Treat* × *Post	**-0.016***	**-0.018****	**-0.025****	**-0.019***	-0.008	-0.011	0.074
	(-1.790)	**(-2.308)**	**(-2.472)**	**(-1.746)**	(-0.796)	(-0.736)	(1.473)
LogEmpNum	-0.004**	-0.002	-0.001	-0.004	0.001	-0.015**	-0.009
	(-2.012)	(-1.286)	(-0.173)	(-1.292)	(0.377)	(-2.298)	(-0.296)
Office FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.398	0.380	0.387	0.384	0.375	0.368	0.406
N. of Obs.	251,359	269,401	148,128	123,659	108,259	41,703	3,283

Table 9: Effect of Bias Audit Law on Job Vacancy Duration

Table 9 examines the effect of the Bias Audit Law on job vacancy duration, using office-quarter observations. Column (1) reports OLS estimates with *LogDuration*—the natural logarithm of the average days to fill a vacancy in a quarter—as the dependent variable. Column (2) reports Poisson estimates with *Duration*—the average days to fill a vacancy in a quarter—as the dependent variable. The primary explanatory variable is the interaction term between *Treat* and *Post*, where *Treat* equals one for offices located in New York City and *Post* equals one for the quarter the law takes effect and thereafter. Standard errors are clustered at the office level. T-statistics are reported below the coefficients. ***, **, and * correspond to 1%, 5% and 10% significance level. Detailed definitions of all variables are provided in Table A.

	(1)	(2)
	OLS Regression	Poisson Regression
DV =	*LogDuration*	*Duration*
Treat* × *Post	**0.202*****	**0.206*****
	(19.638)	**(21.983)**
LogEmpNum	0.061***	-0.040***
	(28.031)	(-20.638)
Office FE	Yes	Yes
Year-Quarter FE	Yes	Yes
R^2 (Pseudo R^2)	0.425	0.289
N. of Obs.	716,495	716,495

Table 10: Effect of Bias Audit Law on Total Hires

Table 10 examines the effect of the Bias Audit Law on total hires, using office-quarter observations. Column (1) reports OLS results where the dependent variable is *LogHiring*, the natural logarithm of the total number of hires in a quarter. Column (2) reports Poisson regression results with *Hiring*—the raw count of hires in a quarter—as the dependent variable. The primary explanatory variable is the interaction term between *Treat* and *Post*, where *Treat* equals one for offices located in New York City and *Post* equals one for the quarter the law takes effect and thereafter. Standard errors are clustered at the office level. T-statistics are reported below the coefficients. ***, **, and * correspond to 1%, 5% and 10% significance level. Detailed definitions of all variables are provided in Table A.

	(1)	(2)
	OLS Regression	Poisson Regression
DV =	*LogHiring*	*Hiring*
Treat* × *Post	-0.006	-0.014
	(-0.897)	(-0.593)
LogEmpNum	0.567***	0.919***
	(88.107)	(128.814)
Office FE	Yes	Yes
Year-Quarter FE	Yes	Yes
R^2(Pseudo R^2)	0.875	0.781
N. of Obs.	556,511	556,511

Table 11: Effect of Bias Audit Law on Firm Performance (Audit Firm Subsample)

This table primarily examines whether the Bias Audit Law affects firms' performance, using only the audit firm subsample. Panel A reports hiring and vacancy-duration results. Panel A reports the effects on hiring diversity and job vacancy duration. Columns (1), (2), (5), and (6) use office-month regressions, while columns (3), (4), (7), and (8) use office-quarter regressions. Columns (1) – (3) estimate the effect on male hiring share (*MaleProp*) and Columns (2) and (4) white hiring share (*WhiteProp*); Columns (5) – (7) estimate the effect on the average number of days to fill a job vacancy (*Duration*) and Columns (6) – (8) estimate the results for the log of average vacancy duration (*LogDuration*). Panel B reports the results of the audit quality. The unit for analysis is at the client-quarter (year) level. Column (1) uses annual regressions with *RestateAnn* (annual restatements) as the dependent variable. Column (2) uses quarter regressions with *RestateQtr* (quarterly restatements) as the dependent variable. The primary explanatory variable is the interaction term between *Treat* and *Post*, where *Treat* equals one for offices located in New York City and *Post* equals one for the month (quarter) the law takes effect and thereafter. Standard errors are clustered at the office level. T-statistics are reported below the coefficients. ***, **, and * correspond to 1%, 5% and 10% significance level. Detailed definitions of all variables are provided in Table A.

Panel A: Employment hiring

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Monthly		Quarterly		Monthly		Quarterly	
	OLS	OLS	OLS	OLS	OLS	Poisson	OLS	Poisson
	MaleProp	*WhiteProp*	*MalePropp*	*WhiteProp*	*LogDuration*	*Duration*	*LogDuration*	*Duration*
Treat × Post	**-0.108*****	0.041	-0.060	0.059	**0.173*****	**0.168***	**0.115****	**0.116****
	(-3.287)	(0.617)	(-1.021)	(1.002)	**(3.948)**	**(4.187)**	**(2.476)**	**(2.572)**
LogEmpNum	0.155**	-0.011	-0.020	-0.002	-0.098	-0.109	0.006	-0.015
	(2.504)	(-0.306)	(-0.722)	(-0.098)	(-1.305)	(-1.340)	(0.457)	(-1.204)
Office FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year-Quarter FE	No	No	Yes	Yes	No	No	Yes	Yes
Year-Month FE	Yes	Yes	No	No	Yes	Yes	No	No
R^2(Pseudo R^2)	0.131	0.265	0.226	0.408	0.310	0.210	0.456	0.266
N. of Obs.	3,847	3,847	2,045	2,045	11,120	11,120	6,161	6,161

Panel B: Audit quality

DV =	(1) *RestateAnn*	(2) *RestateQtr*
Treat* × *Post	**-0.009***	**-0.007**
	(-1.657)	**(-1.265)**
Size	0.001	0.003
	(0.559)	(0.902)
MTB	0.000	0.000
	(0.226)	(0.802)
Lev	0.004	0.008
	(0.505)	(0.537)
ROA	-0.010**	0.005
	(-2.162)	(0.449)
Loss	-0.005	0.002
	(-1.371)	(0.710)
Big4	0.015	-0.018
	(1.490)	(-1.353)
Ln_auditfee	0.004	0.004
	(0.689)	(0.515)
CFO	0.007	0.000
	(0.833)	(1.162)
Auditor_firstyear	0.007	-0.002
	(1.205)	(-0.328)
M&A	-0.000	0.003
	(-0.028)	(0.836)
Num	0.000	0.000
	(0.577)	(0.100)
Foreign	0.016**	0.001
	(2.566)	(0.180)
Disc_ops	0.006	0.002
	(0.825)	(0.322)
Client FE	Yes	Yes
Year-Quarter FE	/	Yes
Year FE	Yes	/
R^2	0.407	0.591
N. of Obs.	15,149	29,732

Table 12: Robustness Tests

This table reports robustness checks for the main estimates of male hiring share and job vacancy duration. Columns (1) use *MaleProp* – the share of male hires – as the dependent variable, while Column (2) uses *LogDuration* – the natural logarithm of the average days to fill a vacancy. Panel A reports results when standard errors are clustered at the metro-area level. Panel B reports results when regressions are at the year-month level. Panel C restricts the sample to offices within New York State. Panel D reports results using firm-quarter fixed effects. Panel D presents results based on an alternative research design, where the control group is composed of offices from companies that have never had an office in New York. The primary explanatory variable is the interaction term between *Treat* and *Post*, where *Treat* equals one for offices located in New York City and *Post* equals one for the quarter (month) the law takes effect and thereafter. T-statistics are reported below the coefficients. ***, **, and * correspond to 1%, 5% and 10% significance level. Detailed definitions of all variables are provided in Table A.

Panel A: Regression at the year-month level

	(1) *MaleProp*	(2) *LogDuration*
Treat* × *Post	**-0.016***** **(-3.634)**	**0.202***** **(20.135)**
LogEmpNum	-0.020*** (-5.493)	-0.002 (-0.918)
Office FE	Yes	Yes
Year-Month FE	Yes	Yes
R^2	0.277	0.348
N. of Obs.	904,018	1,577,516

Panel B: Sample restricted in NY state

DV =	(1) *MaleProp*	(2) *LogDuration*
Treat* × *Post	**-0.014*** **(-1.755)**	**0.165***** **(12.020)**
LogEmpNum	-0.001 (-0.251)	0.101*** (9.386)
Office FE	Yes	Yes
Year-Quarter FE	Yes	Yes
R^2	0.349	0.419
N. of Obs.	31,870	31,749

Panel C: Alternative fixed effects

DV =	(1) *MaleProp*	(2) *LogDuration*
Treat* × *Post	**-0.011***** **(-2.610)**	**0.149***** **(25.490)**
LogEmpNum	-0.000 (-0.920)	0.019*** (41.690)
Firm × Year-Quarter FE	Yes	Yes
R^2	0.219	0.460
N. of Obs.	551,887	711,808

Panel D: Alternative research design

DV =	(1) *MaleProp*	(2) *LogDuration*
Treat* × *Post	**-0.019*****	**0.317*****

	(-2.582)	**(18.032)**
LogEmpNum	0.003	0.121***
	(0.511)	(9.654)
Firm × Year-Quarter FE	Yes	Yes
R^2	0.364	0.425
N. of Obs.	36,254	20,910

Online Appendix Table A: Heterogeneous Effects on Non-White Subgroup Hiring Shares

This table examines how the Bias Audit Law affects hiring shares for different non-white racial groups, using office-quarter observations. Columns (1)–(3) report results for *BlackProp* (proportion of Black hires), *HispanicProp* (proportion of Hispanic hires), and *AsianProp* (proportion of Asian hires), respectively. The primary explanatory variable is the interaction term between *Treat* and *Post*, where *Treat* equals one for offices located in New York City and *Post* equals one for the quarter the law takes effect and thereafter. Standard errors are clustered at the office level. T-statistics are reported below the coefficients. ***, **, and * correspond to 1%, 5% and 10% significance level. Detailed definitions of all variables are provided in Table A.

	(1)	(2)	(3)
DV =	*BlackProp*	*HispanicProp*	*AsianProp*
Treat* × *Post	**0.007****	**-0.008****	0.001
	(2.375)	**(-2.506)**	(0.175)
LogEmpNum	0.000	0.002***	0.000
	(0.076)	(2.813)	(0.497)
Office FE	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes
R^2	0.318	0.374	0.413
N. of Obs.	556,511	556,511	556,511