Workplace discrimination lawsuits against U.S. public corporations *

Casey Dougal, Thomas P. Griffin, Irena Hutton

April 22, 2026

Abstract

We study federal employment-discrimination lawsuits filed against U.S. public companies using a novel dataset that matches Federal Judicial Center filings to CRSP-Compustat and PACER complaint data, covering more than 11,000 firms from 1992 to 2018. Lawsuits concentrate heavily among a small set of repeat defendants, with complaints frequently alleging harassment, retaliation, and systemic misconduct. While industry and geography partly explain lawsuit incidence, persistent firm-level factors account for substantially greater variation. Hiring an external CEO is not associated with a decline in litigation, as post-turnover changes appear linked to the incoming CEO's prior firm litigation history. We find limited evidence of market discipline: stock-price reactions to filings are modest and attenuate as firms accumulate lawsuits, and exogenous shifts in investor base and analyst coverage have little effect on subsequent litigation. These findings suggest that capital-market mechanisms play a minimal role in curbing the persistence of discrimination litigation.

*Casey Dougal: Florida State University, phone: (850) 644-2038, email: cdougal@business.fsu.edu. Thomas Griffin: Villanova University, phone: (610) 519-5460, email: thomas.griffin@villanova.edu. Irena Hutton: Florida State University, phone: (850) 645-1520, email: ihutton@business.fsu.edu. For helpful comments, we thank Vineet Bhagwat, Bernard Black, Dhammika Dharmapala, Daniel Greene, Jillian Grennan, Isaac Hacamo, Michelle Lowry, Lalitha Naveen, Greg Nini, Paige Ouimet, Chris Parsons, Elena Pikulina, Matthew Ringgenberg, Eric Talley, Tracy Yue Wang, Jide Wintoki, Adam Yore, and participants at the University of Washington Fostering Inclusion Workshop (2024), Financial Intermediation Research Society Conference (2023), UNC Kenan-Flagler Finance PhD Alumni Conference (2023), American Law and Economics Association Annual Meeting (2023), European Association of Law and Economics Annual Conference (2023), Financial Management Association Annual Meeting (2023), Conference on Empirical Legal Studies (2022), and Philly Five Conference (2021).

1. Introduction

In *Capitalism and Freedom*, Friedman (1962) argues that competitive markets "protect men from being discriminated against in their economic activities for reasons that are irrelevant to their productivity." Much of the subsequent economic and business discussion adopts the same logic: if discriminatory employment practices are costly, competition and capital markets should discipline them. Yet survey evidence and high-profile workplace-discrimination allegations suggest that workplace discrimination remains common and that employees in large, publicly traded firms continue to report harassment, retaliation, and unequal treatment.[1] This paper asks which firms are publicly accused through litigation, what those accusations look like, and whether markets discipline them.

We address this question by constructing a new dataset of federal civil-rights employment lawsuits filed against U.S. public corporations between 1992 and 2018. Using the Federal Judicial Center (FJC) database, we identify all "Civil Rights–Jobs" cases and algorithmically match corporate defendants to CRSP–Compustat. The resulting FJC–Compustat panel contains more than 94,000 firm-year observations for over 11,000 unique public firms. For a large subsample of cases, we further collect complaint-level information from PACER, which allows us to describe who sues, what conduct is alleged, and how claims are framed in terms of harassment, retaliation, and systemic discrimination.

We organize the analysis around four forces that may shape where and when these lawsuits arise: industry environment, local environment, persistent firm-specific factors, and market discipline. Throughout, we use "culture" as shorthand for persistent norms and

[1]For survey and field evidence that concerns about workplace mistreatment remain salient, see Graham et al. (2022). For evidence that discrimination and harassment can persist even at large employers, see Kline et al. (2022) and Dahl and Knepper (2022). Public controversies surrounding sexual-harassment allegations at listed firms are also consistent with this view; see, for example, Borelli-Kjaer et al. (2021) and Giannetti and Wang (2023).

practices not fully captured by standard economic observables (Bernheim, 1994; Acemoglu and Jackson, 2015; Liu, 2016; Gorton et al., 2022; Graham et al., 2022; Grennan and Li, 2023). The paper first studies how lawsuit incidence varies across industries and places, then asks how much of the remaining variation is tied to persistent firm-specific factors, and finally examines whether investor attention, information production, and stock prices appear to discipline the firms that are publicly accused. Importantly, because the data capture *public, litigated allegations*, the paper speaks most directly to the conditions under which workplace disputes become visible to outside observers and to how markets respond once they do, rather than directly to underlying discrimination.[2]

Three sets of findings emerge. First, federal discrimination lawsuits are highly concentrated. Most public firms are never sued, but a relatively small set of repeat defendants accounts for a disproportionate share of filings. The top 100 firms account for roughly half of all lawsuits while representing less than one-quarter of total employment and less than one-fifth of total assets. The PACER complaints show that the cases are not dominated by narrow technical disputes: harassment appears in 47% of filings, retaliation in 56%, systemic allegations in 42%, and termination in 50%. Almost 79% of plaintiffs report that they raised the issue before filing suit. These patterns suggest that the observed lawsuits often arise from workplace conflicts that are both serious and persistent enough to survive internal escalation and become visible in federal court.

Second, observable litigation varies systematically across industries and geographies, and the relative magnitudes are informative. Industry and economic-area fixed effects each ex-

[2]Many workplace conflicts are resolved internally or through administrative channels and never reach federal court; others may be deterred by the threat of retaliation or weak outside options (Hersch, 2011, 2018; Dahl and Knepper, 2022). Conversely, some federal suits are dismissed or resolved in favor of the defendant. Accordingly, the paper is best viewed as studying the incidence and consequences of *observable discrimination allegations* rather than adjudicated wrongdoing or the total amount of discrimination inside firms.

plain roughly 3% of the variation in whether a firm is sued in a given year, and jointly they explain about 6%, indicating that broad sector and location both matter. In peer-rate regressions, a one-percentage-point increase in local non-industry lawsuit rates predicts about a 1.1% increase in a firm's own expected lawsuits, while a comparable increase in non-local same-industry lawsuit rates predicts about a 2.5% increase. Read against Parsons et al. (2018), these patterns suggest that the employment setting has a stronger national-industry component than the financial misconduct setting, even though local environment still matters materially.

The strongest persistence, however, appears at the firm level. Variance decompositions show that persistent firm effects account for roughly 56% of the total variation in lawsuits, compared with about 7% for CEO effects, although CEO influence remains economically meaningful. There is no evidence to suggest that hiring an external CEO decreases litigation; if anything, there is marginally significant evidence that it increases litigation. To understand that result, we examine heterogeneity based on the litigation history of the incoming CEO's prior firm and find that lawsuit incidence rises more after CEOs arrive from high-litigation firms than after CEOs arrive from low-litigation firms. Taken together, these results suggest that repeat-litigation status is not simply an artifact of industry or geography. It is primarily a firm-level phenomenon, with management helping shape how that persistence unfolds.

Third, we find only limited evidence that outside markets materially discipline these patterns once they become observable. Leadership transitions provide some evidence of change, but they do not produce the broad reset in litigation one would expect if external discipline routinely overturned firms' underlying litigation profiles. We therefore turn to two cleaner market-based shocks. A Russell 1000/2000 regression-discontinuity design generates plausibly exogenous changes in index assignment, and thus in investor base, around the annual cutoff, yet those shifts have no detectable effect on lawsuit incidence. Brokerage closures and mergers likewise generate quasi-exogenous reductions in analyst coverage, but lower analyst following does little to affect subsequent lawsuits. Shareholders do react when lawsuits are filed, but the pricing response is modest and front-loaded: the first suit generates the most negative abnormal return, while additional suits become progressively less informative. Markets appear to learn from public allegations, but not in a way that eliminates persistent repeat defendants.

The paper's primary contribution is descriptive. To our knowledge, it provides the first comprehensive panel of federal employment-discrimination lawsuits against U.S. public corporations and uses it to document where these allegations arise, how concentrated they are, and what kinds of claims they contain. In that respect, the paper sits between work on discrimination measured in worker-level or administrative settings and work that studies litigation, scandals, and other public signals of firm misconduct (Bizjak and Coles, 1995; Gande and Lewis, 2009; Gormley and Matsa, 2011; Wang et al., 2010; Karpoff et al., 2017). The cases are tied directly to workplace discrimination, but they are observed only when disputes become public through federal court. That feature limits causal interpretation, but it also makes the data well suited to studying what outside observers actually learn from these allegations.

More broadly, the paper studies the same basic question as Parsons et al. (2018) in a different setting: when public allegations cluster, how much of that clustering is associated with industry, place, and persistent firm-specific factors? In our setting, those margins plausibly reflect a different mix of employment practices, operating models, local institutions, and reporting incentives than in the financial misconduct setting.

Taken together, the evidence points in one direction. Public discrimination litigation is not spread evenly across public firms, nor is it strongly disciplined by markets once it becomes part of a firm's history. These cases are concentrated in particular industries, local environments, and especially firms. We interpret this pattern as evidence of persistent heterogeneity in the conditions that give rise to federal discrimination allegations, while remaining agnostic about the precise mix of underlying conduct, reporting, settlement, and legal selection that maps those conditions into observed filings.

The rest of the paper proceeds as follows. Section 2 describes the legal setting, data construction, and basic descriptive patterns. Section 3 studies industry environment using fixed effects and peer lawsuit rates and compares those patterns to the geography-of-misconduct results in Parsons et al. (2018). Section 4 examines local environment and cultural proxies. Section 5 analyzes persistent firm-specific and CEO-specific components of lawsuit incidence, including a separate external-turnover analysis of CEO transition types. Section 6 presents three distinct tests of outside discipline: index inclusion, analyst-coverage shocks, and stock-price reactions, and then draws the broader conclusion from those experiments.

2. Institutional Background and Data

Federal law prohibits discrimination in employment based on race, color, religion, sex, national origin, age, and disability, among other protected characteristics. Title VII of the Civil Rights Act, the Age Discrimination in Employment Act, and the Americans with Disabilities Act provide the principal federal framework under which employees can challenge discriminatory treatment. Before filing suit in federal court, claimants typically must file a charge with the Equal Employment Opportunity Commission (EEOC) or an equivalent state agency. If the agency issues a "right-to-sue" letter, employees may then bring a civil action in federal district court, where cases can culminate in settlement, dismissal, or judgment. Prior work uses these legal processes to study the interaction between civil-rights enforcement, workplace policies, and firm outcomes (Hersch, 1991; Kalev and Dobbin, 2006; Knight et al., 2022; Kline et al., 2022).

2.1. FJC–Compustat panel and PACER complaints

We construct our main measure of discrimination activity from the FJC database of civil filings in U.S. district courts. We identify all cases coded as “Civil Rights–Jobs” between 1992 and 2018 and match corporate defendants to CRSP–Compustat using standard name- and CUSIP-based algorithms. The resulting FJC–Compustat panel contains 94,113 firm-year observations for 11,331 unique U.S. public corporations. For each firm-year, we count the number of new discrimination suits filed against the firm and merge in firm characteristics, headquarters location, and Fama–French 12-industry classification. This structure allows us to study both the incidence of lawsuits and their relation to firm, industry, and geographic attributes in the spirit of Parsons et al. (2018) and related work on litigation-based measures of misconduct (Karpoff et al., 2017; Gorton et al., 2022; Graham et al., 2022).

To better understand what these filings contain, we complement the FJC panel with a detailed subsample of complaints from PACER. For a large subset of corporate defendants, we download the complaints and construct case-level variables describing plaintiffs' characteristics, the alleged conduct, and requested remedies. At the firm level, the PACER subsample covers 21,127 firm-year observations for 1,681 public companies between 1999 and 2018, with coverage tilted toward larger and more visible firms.

The PACER variables are extracted from unstructured complaint text using a structured prompt applied case by case to the complaint document. The extraction protocol asks for docket and filing metadata; observable plaintiff characteristics such as gender, race, age, job title, tenure, compensation, and employment status; the protected classes and forms of discrimination alleged; short summaries of the underlying conduct; the identities of accused individuals; internal reporting and grievance activity; indicators of comparative treatment, quoted discriminatory language, and broader workplace patterns; and the remedies sought. When the complaint does not state a field, the protocol records it as "Not Mentioned," which helps distinguish true absence from non-disclosure. Appendix A.1 reproduces the prompt used to generate these structured PACER fields.

Table 1 summarizes the PACER cases by protected class. Several patterns help clarify what our lawsuit measure captures. The typical filer is a mid-career worker, consistent with the idea that many disputes arise after employees have spent meaningful time within the firm. Harassment and retaliation are pervasive across protected classes, indicating that a large share of suits involve hostile work environments and adverse responses to reporting. Systemic claims referring to policies or practices affecting groups of employees are also common, especially in gender- and race-related complaints, and many cases involve multiple alleged bases of discrimination. Overall, the PACER evidence suggests that the observed federal filings often reflect serious workplace conflicts rather than purely technical legal disputes.

Several features of Table 1 deserve emphasis because they shape how the later tests should be interpreted. The median plaintiff has five years of experience at the firm and is about 49 years old, suggesting that these are not predominantly disputes involving very short-lived job matches. Executives account for only about 10% of cases, managers for 26%, and rank-and-file employees for 64%, which helps explain why local and industry operating environments matter so much in later sections: the modal case concerns ordinary employment relationships rather than executive bargaining. The allegation mix is also revealing. Termination appears in 50% of cases, retaliation in 56%, and harassment in 47%, while only 3% of filings concern hiring. This composition suggests that the dataset is more informative about breakdowns in ongoing employment relationships than about discrimination at the initial hiring margin. Finally, almost 57% of cases request reinstatement or back pay, and almost 79% of plaintiffs report prior internal reporting, both consistent with disputes that were first contested within the organization before escalating outside it.

At the same time, Table 1 should not be read as converting allegations into verified wrongdoing. Complaints are one-sided legal documents, and the set of cases that reaches federal court is shaped by employees' incentives, attorney supply, internal dispute resolution, and expected litigation outcomes. We therefore use the PACER data primarily to characterize the nature of the allegations that become public, not to infer that every claim reflects a realized violation.

2.2. Distribution of discrimination litigation

Table 2 documents how discrimination lawsuits are distributed across firms in the FJC–Compustat panel. The mean number of suits per firm-year is 0.39, but the distribution is highly skewed. Approximately 73% of firms are never sued during the sample period; about 11% are sued in exactly one year; roughly 10% are sued in two to five years; and only about 3% are sued in six to ten years. Lawsuits are therefore rare for most public firms, but a meaningful minority become repeat defendants.

The concentration is especially pronounced among large employers, but size alone does not explain it. Firms in the top employee quartile account for 87% of all lawsuits, and firms in the top asset quartile account for nearly 80%. Yet even within the upper tail, lawsuits are concentrated among a subset of repeat defendants. The 100 firms with the most discrimination suits account for roughly 49% of all lawsuits in the sample, but only about 24% of total employment and 19% of total assets. In other words, large scale is clearly part of the story, but it is not the entire story: some very large employers are sued far more often than others. This concentration mirrors patterns documented for other forms of litigation and misconduct (Biggerstaff et al., 2015; Gormley and Matsa, 2011; Parsons et al., 2018), and it foreshadows one of the paper's main themes: public discrimination allegations are not randomly scattered across firms.

3. Industry Culture

We begin with industry environment. Industries differ in their technologies, customer interactions, workforce composition, and supervisory structures, all of which may affect both the prevalence of workplace conflict and the likelihood that disputes escalate into public litigation. We refer to these persistent industry-level differences as "industry culture," but the broader point is simply that firms operating in different sectors face systematically different conditions under which employment disputes arise and are litigated.

Table 3 reports R^2 values and F-statistics from linear probability models in which the dependent variable is an indicator for whether a firm faces at least one discrimination lawsuit in a given year. Year fixed effects alone explain very little variation ($R^2 \approx 0.0025$). By contrast, Fama–French 12-industry fixed effects on their own raise the R^2 to roughly 0.031, and economic-area fixed effects deliver a nearly identical fit. When year, industry, and area effects are included together, the R^2 rises to about 0.06. The F-tests show that industry effects are strongly significant on their own and after conditioning on year, while area effects remain strongly significant after conditioning on both year and industry.

These magnitudes are modest in absolute terms but informative in context. Lawsuits are rare events, so a single-digit R^2 can still carry meaningful explanatory content. Moving from year effects alone to industry or area effects increases explanatory power by more than an order of magnitude. The fixed-effects evidence points to systematic sectoral differences in the environments that generate observable discrimination litigation. It also provides a benchmark for the later firm-effects results: industry and geography matter, but they leave substantial within-industry and within-location heterogeneity unexplained.

These results are also useful to read alongside Parsons et al. (2018). In that paper's baseline geography tables, Parsons, Sulaeman, and Titman show that misconduct is not evenly distributed across cities and that geographic fixed effects contain meaningful explanatory power. Our Table 3 delivers the same broad message in a different empirical setting: location matters, but so does industry. Relative to Parsons et al. (2018), the employment-litigation setting appears somewhat less purely geographic and somewhat more jointly shaped by sector and place. That contrast is economically intuitive. Many employment disputes arise from line management, workforce organization, and operating models that are partly industry-specific, whereas the financial misconduct setting in Parsons et al. is naturally more tied to local executive and professional networks.

Table 4 takes a more granular approach by relating a firm's lawsuit incidence to the lagged lawsuit rates of its peers. We estimate Poisson regressions of the number of lawsuits a firm faces in year t on three peer variables measured in $t-1$: the lawsuit rate among local non-industry peers in the same economic area, the lawsuit rate among local same-industry peers, and the lawsuit rate among non-local same-industry peers outside the firm's area. All specifications control for lagged firm characteristics and year effects.

Two patterns stand out. First, firms are more likely to be sued when other employers in the same local area are more frequently sued, even if those employers are in different industries. In column (1), a one-percentage-point increase in local non-industry lawsuit rates predicts roughly a 1.1% increase in expected lawsuits at the focal firm; for local lawsuits in column (2), the corresponding semi-elasticity rises to about 2.3%. This pattern is exactly what one would expect if local legal infrastructure, reporting norms, or dispute-escalation processes matter. Second, firms are also more likely to be sued when their national industry peers outside the local area have higher lawsuit rates. In column (1), the coefficient on non-local same-industry peers is 0.025, more than twice as large as the local non-industry effect, and in column (3) it is 0.028 for non-local lawsuits. That pattern is hard to reconcile with a purely local litigation story and instead points toward industry-wide operating environments or employment models.

Column (2) also sharpens the role of local same-industry peers. That coefficient is 0.005 and statistically significant, whereas it is otherwise generally small and insignificant, with only a marginally significant coefficient in the PACER economic-lawsuit specification. Put differently, once the dependent variable is restricted to *local* lawsuits, both parts of the local environment matter: the broad area portfolio and the narrower same-industry local portfolio. The magnitudes are informative. The coefficient on the broad local portfolio in column (2) is still much larger than the same-industry local coefficient, so the dominant local force appears to be the area-wide dispute environment rather than a narrowly defined cluster of nearby industry peers. But the same-industry local effect does contain incremental information for this local outcome, which is exactly what one would expect if geographically proximate firms in the same sector share supervisors, labor pools, employment practices, and plaintiffs' attorneys that are especially relevant for disputes filed close to headquarters.

The broader implication is that the relevant comparison set depends on the outcome being studied: for local lawsuits, it is local in both a geographic and industry sense; for broader lawsuit incidence, it is better summarized by a combination of the general local environment and the firm's national industry environment.

The PACER subsample points in the same direction. The local non-industry and non-local same-industry effects are strongest for non-economic lawsuits and for employee lawsuits, whereas the coefficients are small and insignificant for executive cases. Harassment, retaliation, and day-to-day workplace conflict are more likely than executive disputes to be shaped by sectoral labor practices and local employment norms. The clustering is concentrated where one would expect it if routine workplace interactions are central to the litigation process.

We do not interpret these peer coefficients as causal spillovers. They may reflect correlated exposure to unobserved sectoral or regional shocks, common legal environments, or plaintiffs' attorney activity. Even so, the fixed-effects and peer-regression evidence together paint a clear picture: discrimination lawsuits cluster by industry in ways that cannot be reduced to simple time trends alone. Sector-specific operating environments appear to be an important part of the backdrop against which federal discrimination allegations become observable.

The comparison with Parsons et al. (2018) is again instructive. Their peer-effect results emphasize that nearby firm misconduct is strongly associated with the focal firm's misconduct, a pattern they interpret as consistent with local norms and peer effects. Our Table 4 points in the same general direction, but the relative weights differ. Local non-industry lawsuit rates matter in our data, yet the non-local same-industry effect is at least as important and often larger. Relative to Parsons et al., the employment setting therefore appears to have a stronger industry-wide component. That difference is substantively useful because it suggests that the relevant transmission mechanism is not just local social spillovers; it also runs through national industry practices and organizational templates.

4. Local Culture

We next turn to local environment. A large literature in economics and finance shows that local prejudice, social norms, media markets, and institutional conditions shape economic behavior and corporate outcomes (Charles and Guryan, 2008; Kumar et al., 2011; Liu, 2016; Lins et al., 2017; Hayes et al., 2021; Heese et al., 2022). In our setting, local conditions may matter through several channels at once: the incidence of underlying workplace problems, workers' willingness to report them, the supply of plaintiffs' attorneys, and the way judges and juries treat such claims. Our goal in this section is not to isolate a single mechanism, but to document whether observable federal litigation is systematically related to local attitudes and environments.

The fixed-effects evidence in Table 3 already suggests that geography matters nearly as much as industry. Table 4 reinforces that message: local non-industry lawsuit rates strongly predict a firm's own lawsuits even after controlling for national same-industry rates. We next examine whether these patterns line up with direct measures of local demographics and cultural proxies.

Table 5 reports Poisson regressions that relate the number of discrimination lawsuits to lagged firm, industry, and headquarters-area characteristics. Several local variables are especially informative for *local* lawsuits, that is, suits filed in the firm's own headquarter area. Firms headquartered in richer areas face fewer local suits, firms in areas with a larger

Black population share face more local suits, and firms headquartered in rural areas face fewer local suits. These correlations are consistent with local differences in labor markets, legal access, and reporting environments.

Our two main cultural proxies point in the same direction. The headquarters racial-bias index enters positively and significantly in the local-lawsuit regressions, with a coefficient of 0.175 in column (2), while the indicator for "least sexist" states enters negatively and strongly significantly, with a coefficient of -1.327. In economic terms, these are among the largest local correlates in the table. Their magnitudes also matter relative to the other controls. For example, the sexism coefficient is substantially larger in absolute value than the coefficient on log income per capita, suggesting that the cultural proxy is not just standing in for local affluence.

The cross-column patterns are just as important as the individual coefficients. These variables matter most for local lawsuits, much less for non-local lawsuits, and only selectively in the PACER subsample. That is exactly the pattern one would expect if the proxies are picking up something genuinely local rather than a generic omitted firm trait. Within the PACER subsample, the heterogeneity is more mixed than a single summary would suggest. The sexism proxy loads more clearly in non-economic and non-executive cases, but the racial-bias proxy is not confined to those disputes and is also significant for executive lawsuits. The safest conclusion is therefore not that local norms matter only for routine workplace interactions, but that local environment remains relevant within the PACER sample while operating somewhat differently across the racism and sexism proxies.

These results should be interpreted carefully. A positive coefficient on local racial bias, or a negative coefficient on local gender egalitarianism, does not by itself identify whether the relevant margin is more discrimination, more willingness to litigate, less internal resolution, different settlement behavior, or a different judicial environment. For that reason, we view the local proxies as capturing a broad local ecosystem rather than a single behavioral channel. Still, the evidence is difficult to attribute entirely to chance: observable federal discrimination litigation is substantially more common in some local environments than in others, even after conditioning on firm characteristics and industry.

Table 6 adds another perspective by interacting the local proxies with an indicator for geographically concentrated operations. The racial-bias measure matters most for firms whose operations are concentrated in relatively few states, suggesting that headquarter environment is more closely connected to observed litigation when the firm is less geographically diffuse. For the sexism proxy, the "least sexist state" indicator is negative regardless of geographic concentration, and the interaction is small. This difference in patterns cautions against overly literal interpretation, but it is consistent with the idea that some local influences operate most strongly when firms are tightly embedded in a particular place.

The magnitudes in Table 6 are informative. The stand-alone "most racist state" indicator is imprecisely estimated, but its interaction with concentrated operations is positive and statistically significant across nearly all specifications, ranging from about 0.55 to 0.95. This means that the relevant margin is not simply being headquartered in a high-bias state; it is being both headquartered there and operationally tied to that environment. For employee lawsuits in column (7), for example, the interaction coefficient is 0.926. By contrast, the "least sexist state" effect is already large and negative for local lawsuits without much interaction effect. One way to summarize the table is that racial-bias proxies appear to matter through local embeddedness, while the sexism proxy appears to capture a broader state-level environment that matters even for firms that are more geographically dispersed.

The local results make two points. First, geography matters even after rich firm controls and industry fixed effects are included. Second, the most plausible interpretation is not a single clean channel, but a local package of norms, institutions, and reporting conditions. That interpretation is less sharp than a causal estimate of discrimination per se, but it is also more credible given what the data actually observe.

Overall, the local evidence complements the industry evidence. Observable discrimination lawsuits are not just concentrated in certain sectors; they are also concentrated in certain places. Geography appears to matter through a bundle of local norms and institutions that shape which disputes become visible in federal court.

5. Firm Culture

Industry and geography explain meaningful variation in lawsuit incidence, but substantial heterogeneity remains across firms operating in the same broad environments. We now turn to persistent firm-specific factors. We use "firm culture" as a convenient label, but the evidence more generally speaks to durable differences in firms' employment practices, internal governance, dispute-resolution systems, and tolerance for behavior that can culminate in public litigation.

Table 5 shows that observable firm economics explain some, but far from all, of the variation in lawsuits. Size is by far the strongest correlate: the coefficient on log employees is 1.806 in the total-lawsuit specification, and both employment and assets per employee are positive and highly significant across all columns. Workforce scale is therefore the clearest mechanical predictor of litigation. Beyond size, the coefficients on other firm fundamentals are less uniform and do not point to a straightforward economic story. Sales growth is negative in most specifications, leverage is positive in several columns, market-to-book loads positively in some cases, and the ROA coefficients are not stable across outcomes. Taken together, Table 5 suggests that standard firm observables matter, but not always in intuitive ways, and a large share of the heterogeneity in litigation lies beyond what those observables can explain.

The natural next question is whether that remaining variation reflects persistent firm-specific differences or instead the influence of individual managers. To quantify the persistent firm-specific component more directly, Table 7 reports AKM-style variance decompositions for seven lawsuit outcomes, including total lawsuits, local and non-local lawsuits, economic and non-economic lawsuits, and executive and employee lawsuits. We estimate linear models with lagged firm controls, firm fixed effects, and CEO fixed effects, then decompose the variance of each outcome into the shares attributable to observables, firm effects, CEO effects, and the residual.

Firm effects dominate by a wide margin. Controls alone explain 22.9% of the variance in total lawsuits and only 3.9% of the variance in local lawsuits. Once firm and CEO fixed effects are added, the firm component accounts for 55.7% of total-lawsuit variance, 27.7% of local-lawsuit variance, and 41.4% of employee-lawsuit variance. CEO effects are present but materially smaller, ranging from about 6% to 9% across outcomes. The comparison is especially revealing for the core outcome: for total lawsuits, the firm component is roughly eight times as large as the CEO component. Even where CEO effects are relatively larger, such as local lawsuits, they remain far below the firm effect.

We interpret this finding as evidence of persistent heterogeneity across firms, not as proof that the fixed effect is "pure culture." A firm fixed effect can absorb a broad bundle of time-invariant features, including HR systems, promotion structures, workforce composition, and the way complaints are handled internally. That said, the magnitude of the firm component is economically important. Whatever the exact underlying mechanisms, some firms are persistently much more likely than others to generate public federal discrimination litigation, and that persistence is much stronger than what can be explained by observable economics or by the average CEO effect.

The bottom panel of Table 7 also argues for caution in interpreting the CEO effects. The firm–CEO connected set is small: only about one-third of one percent of the estimation sample is in the largest connected component, and there are relatively few CEO movers. As a result, the decomposition is more informative about persistent differences across firms than about precisely estimated manager-specific effects. That is sufficient for our purposes. The question is whether litigation is better characterized as a firm-level or manager-level phenomenon, and the decomposition clearly points to the former, even if the exact CEO share is estimated with noise. Because the connected-set limitations make it difficult to pin down CEO effects precisely, we examine the CEO question using a different design.

Table 8 addresses the CEO question from a complementary angle. Rather than decomposing variance into firm and CEO components, the table asks whether lawsuit incidence changes around external leadership transitions and whether those changes depend on the discrimination status of the CEO's origin and destination firms. The setup is motivated by the manager-effects literature, especially Bertrand and Schoar (2003) and Fee et al. (2013), which uses executive moves to separate persistent firm traits from managerial style. The same logic applies here. If discrimination litigation is primarily a firm-level phenomenon, ordinary turnover should have limited effects on lawsuits. If CEOs carry employment practices across firms, the transition type should matter as well.

The table implements this idea with stacked event-time difference-in-differences specifications centered on external CEO turnover events. Panel A clarifies how much identifying variation is available once turnover events are classified by the discrimination status of the origin and destination firms. There are 367 external CEO moves in total, but only 167 can be classified. Among those classified moves, 30 are high-to-low, 57 are low-to-high, 42 are high-to-high, and 38 are low-to-low. The key point is the limited size of the classified cells. That fact should temper how strongly one reads the more disaggregated estimates in Panel B.

Panel B then traces lawsuit incidence in the three years before and after each turnover, using firms that have not yet experienced a turnover in the same calendar window as controls and including firm-cohort and year-cohort fixed effects. Columns (1) and (2) report the average effect across all external CEO turnovers. There is no evidence to suggest that hiring an external CEO decreases litigation; if anything, there is marginally significant evidence that it increases litigation. The coefficient on *Post CEO Turnover* is 0.147 without controls and 0.144 with controls, which in a PPML model corresponds to increases of roughly 15–16% in expected lawsuits. To understand this result, columns (3) and (4) explore heterogeneity based on the litigation history of the incoming CEO's prior firm.

Columns (3) and (4) show that the average effect masks very different post-turnover paths across transition types. The baseline *Post CEO Turnover* coefficient is negative in these columns, implying that for the omitted category of other external-turnover events, expected lawsuits fall by about 31–33% after the turnover. The high-to-low interaction is then extremely large: coefficients of 3.196 and 3.079 imply very sharp post-turnover increases relative to the omitted category. Taken literally, the implied magnitudes are unusually large, so they should be interpreted with caution given the small number of classified high-to-low transitions in Panel A and the endogenous nature of CEO matching. We therefore view these coefficients primarily as evidence of strong heterogeneity across transition types rather than as precise causal effect sizes. By contrast, the low-to-high interaction is negative, with coefficients of -0.471 and -0.377, implying post-turnover declines of roughly 53–58% in expected lawsuits once combined with the negative baseline coefficients. Economically, the table therefore does not point to a simple story in which outside CEOs either raise or lower litigation on average. Instead, it suggests that what matters is the type of origin-to-destination transition.

The interpretation remains cautious. CEO hiring is endogenous, and the classified transition cells in Panel A are small, so the estimates should not be read as clean causal CEO effects. Even so, the table disciplines the broader narrative. The data do not support a view in which leadership is irrelevant, but they also do not suggest that average CEO turnover reliably resets a firm's litigation regime. The most natural reading is hierarchical: firm-specific persistence is first-order, while CEO background and transition type matter at the margin.

This section points to a clear ranking. Standard observables help explain which firms are sued, but persistent firm-specific differences explain far more. The CEO evidence is not negligible: the AKM decomposition assigns a smaller but nonzero role to CEOs, and the external-turnover results show that litigation can change across transition types. Thus, we find firms matter most, but CEOs matter too.

6. Market Discipline

Our final question is whether outside markets appear to discipline firms that face discrimination allegations. Becker's framework suggests that competition can make discrimination costly (Becker, 1957; Hart, 1983). In our setting, any such discipline would have to operate on *observable allegations*: investor scrutiny, analyst attention, or stock-price penalties could induce firms to reduce the conduct, reporting failures, or dispute escalation that result in public lawsuits. We therefore study whether plausibly exogenous changes in investor attention and information production affect lawsuit incidence, and how strongly shareholders react when a lawsuit is filed.

Index inclusion and exogenous shifts in investor base

Table 10 implements a local-linear regression discontinuity design around the Russell 1000/2000 cutoff. Crossing the cutoff generates a sharp change in index membership and therefore in ownership structure and investor attention. The design follows Gantchev et al. (2022) in its implementation with the inclusion of the improved assignment procedure in Ben-David et al. (2019). Firms near the threshold are similar in size, but they experience discontinuous changes in index assignment that alter the composition of their shareholder base for reasons unrelated to their underlying employment litigation risk.

The first-stage estimates confirm that the design has substantial mechanical power. Falling on the Russell 2000 side of the cutoff raises actual July Russell 2000 membership by 0.35 to 0.58 across the reported bandwidths, and the first-stage F-statistics range from 11.8 to 85.7. In other words, the discontinuity generates a meaningful and precisely estimated shift in investor base. This matters because any second-stage null is only informative if the instrument actually moves index assignment in a material way; here it clearly does.

The second stage contains the main result. The coefficients on Russell 2000 membership in the lawsuit equation are -0.099, -0.020, and 0.015 across the three bandwidths. Because the dependent variable is $\log(1+\#\text{Lawsuits})$, these coefficients are not directly interpretable as percentage changes in lawsuits themselves. The more important point is that the estimates are statistically weak, unstable in sign across bandwidths, and centered near zero. In particular, the narrow-bandwidth estimate is imprecise enough that the data cannot rule out economically meaningful effects in either direction. Taken together, the index-inclusion design does not provide clear evidence that exogenous shifts in investor base materially affect future lawsuit incidence.

That null result matters because it speaks directly to a common monitoring hypothesis. If changes in investor composition—especially the shift toward more indexed ownership and the accompanying increase in visibility—materially discipline firms on employment-related conduct, those effects should be visible in a design of this kind. Table 10 instead suggests that even plausibly exogenous changes in investor base are not enough to produce detectable changes in the rate at which employment disputes escalate into federal court.

Brokerage closures, analyst attention, and lawsuit incidence

Table 11 turns to analyst coverage using brokerage closures and mergers. The design follows the analyst-coverage shock literature that uses brokerage events as quasi-exogenous reductions in coverage, beginning with Kelly and Ljungqvist (2012) and extending to later work on analyst monitoring and corporate policy. Here we follow the approach of Guo et al. (2019). The identifying idea is that when a brokerage closes or merges, some firms lose coverage for reasons unrelated to their contemporaneous fundamentals. If analysts are an important monitoring channel for employment-related misconduct, these shocks should affect lawsuit incidence as well as coverage itself.

The first result in the table is that the design moves analyst attention in the intended direction. Exposure to future closures significantly reduces subsequent analyst coverage: the coefficient of -0.071 in column (1) implies that a firm whose entire analyst roster is exposed would experience about a 6.9% decline in expected coverage, and even a 10-percentage-point increase in exposure implies a decline of roughly 0.7%. That is not a massive first stage, but it is precisely estimated and confirms that the closure events generate a real deterioration in outside information production.

The second and third columns are the substantive result. In column (2), the coefficient on exposure in the future-lawsuit regression is -0.041, implying a small decline in expected lawsuits, but the estimate is economically modest and statistically weak. Column (3) reports the corresponding pre-trend regression using lawsuits from years $t-3$ to $t-1$. The coefficient is also negative and similarly imprecise. That pattern matters because it suggests that the weak future-lawsuit estimate is not a sharp post-shock response, but instead sits against a background of similarly noisy pre-existing differences. Put differently, once the design is used to isolate plausibly exogenous changes in analyst attention, there is little evidence that lower coverage translates into a clear increase or decrease in observed discrimination litigation.

This result is informative even if it is not dramatic. Analysts matter for price discovery, external scrutiny, and managerial visibility in many settings. Yet Table 11 suggests that the processes generating employment litigation are not especially elastic to moderate changes in sell-side coverage. One plausible interpretation is that analyst scrutiny matters more for how the market interprets lawsuits once they are filed than for whether workplace disputes arise, escalate, and become public in federal court in the first place.

Stock-price reactions and the information content of lawsuits

Tables 8, 10, and 11 suggest that the market-discipline evidence is limited. External CEO transitions can be followed by meaningful changes in litigation for particular origin-to-destination matches, but neither ordinary turnover nor plausibly exogenous shifts in investor base and analyst attention produce the kind of broad response one would expect if external discipline were central to lawsuit incidence. The through-line is persistence: public employment litigation appears tied to relatively durable features of firms and their environments. We turn next to stock-price reactions to lawsuit filings, which provide the most direct evidence on whether markets respond when these allegations become public.

Table 12 shows that discrimination lawsuits do elicit a market response, but the magnitude is modest. In the full sample, the mean five-day CAR is -0.041% using the market-adjusted benchmark and -0.059% using the Fama–French benchmark, or about 4 to 6 basis points. In dollar terms, the mean CAR is about -25.7 million using the market-adjusted benchmark and -23.3 million using the Fama–French benchmark, while the median dollar CAR is about -1.1 million under either benchmark. That gap between means and medians makes clear that the dollar effects are highly skewed by firm size. For the typical lawsuit, the valuation effect looks modest and may be closer to direct legal and settlement costs than to a large reputational repricing, even though the implied dollar losses can still be substantial for the largest firms in the sample.

The defendant-history splits in Table 12 are especially revealing. In Panel B, a firm's first lawsuit is associated with a five-day CAR of about -0.15% at the mean and -0.24% at the median, compared with -0.002% and -0.073%, respectively, for firms that have already accumulated eleven or more suits. Put differently, the first allegation against a firm generates a noticeably more negative market response than a suit filed against a seasoned repeat defendant. The difference is about 15 basis points in the means and 17 basis points in the medians. For a large public company, that gap can easily amount to tens of millions of dollars in incremental market-value loss. This is exactly the pattern one would expect if the first few lawsuits teach the market something about a firm's workplace environment, while later suits mostly confirm what investors already know. In that sense, the market learns, but the pricing schedule is front-loaded: the first suit carries the most information, and additional suits add comparatively little.

Taken together, the market evidence points to limited rather than nonexistent discipline. Investors do react to discrimination lawsuits, and the first allegation against a firm carries negative information. But the pricing response is modest, and exogenous shifts in investor base or analyst attention do not materially alter lawsuit incidence in our tests. Markets appear capable of incorporating this information into valuation, yet less capable of eliminating persistent high-litigation firms.

7. Conclusion

We study federal employment-discrimination lawsuits against U.S. public corporations. Using a new panel that links Federal Judicial Center "Civil Rights–Jobs" filings to CRSP–Compustat, together with complaint-level information from PACER, we characterize which firms are publicly accused in federal court, what those allegations contain, and whether outside markets discipline firms that are repeatedly sued.

The paper is deliberately descriptive. The data capture *public, litigated allegations*, not the full universe of workplace discrimination, and not only cases in which wrongdoing is ultimately established. Observed litigation reflects some combination of underlying conduct, reporting incentives, internal dispute resolution, attorney supply, settlement, and judicial process. Within those limits, the empirical patterns are sharp. Federal discrimination lawsuits are highly concentrated among a small set of repeat defendants. The PACER complaints indicate that these cases commonly involve harassment, retaliation, and systemic allegations, rather than narrow technical disputes. Industry and geography both matter, but the most persistent source of heterogeneity lies at the firm level.

That ranking is central to the paper's interpretation. Standard observables, including size, growth, leverage, and profitability, explain part of the variation in lawsuit incidence, but not always in intuitive ways, and persistent firm effects explain far more. CEO effects are present, and the turnover evidence suggests that managerial background can matter at the margin. Even so, the dominant fact in the data is firm persistence: some firms are repeatedly exposed to public discrimination litigation in ways that are not well explained by industry, geography, or ordinary leadership change alone.

The evidence on outside discipline is correspondingly limited. Exogenous shifts in investor base do not alter lawsuit incidence, and reductions in analyst coverage have little detectable effect on whether disputes surface in federal court. Shareholders do react to lawsuit filings, especially when a firm is accused for the first time, but those pricing effects are modest and attenuate quickly for repeat defendants. Markets appear to learn from these allegations without imposing penalties large enough to erase the strong persistence with which certain firms are sued.

More broadly, the paper opens a new empirical window on workplace-discrimination allegations at public firms. Because the data identify which disputes become visible in federal court, they provide a benchmark for studying how internal employment practices, legal institutions, public scrutiny, and capital markets interact. That benchmark suggests a natural agenda for future work: linking allegations more directly to internal governance choices, human-capital policies, and labor-market outcomes within firms; studying which institutional margins determine whether workplace disputes remain internal, proceed through administrative channels, settle quietly, or become public federal cases; and examining how investors, workers, and firms learn from repeated allegations, and whether that learning changes hiring, retention, financing, or organizational structure.

For boards, investors, and policymakers, the evidence is clear on one point: market forces alone are unlikely to eliminate firms that repeatedly generate these disputes. Persistent high-litigation status is not erased by ordinary leadership turnover, changes in investor composition, or shifts in analyst coverage. Understanding what does alter that persistence, and at what cost, remains an important task for future research.

References

Acemoglu, D., Jackson, M. O., 2015. History, expectations, and leadership in the evolution of social norms. Review of Economic Studies 82, 423–456.

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

Ben-David, I., Franzoni, F., Moussawi, R., 2019. A note to "do etfs increase volatility?": An improved method to predict assignment of stocks into russell indexes. Journal of Finance Replications and Corrigenda Web-only publication.

Bernheim, B. D., 1994. A theory of conformity. Journal of Political Economy 102, 841–877.

Bertrand, M., Schoar, A., 2003. Managing with style: The effect of managers on firm policies. Quarterly Journal of Economics 118, 1169–1208.

Biggerstaff, L., Cicero, D. C., Puckett, A., 2015. Suspect CEOs, unethical culture, and corporate misbehavior. Journal of Financial Economics 117, 98–121.

Bizjak, J. M., Coles, J. L., 1995. The effect of private antitrust litigation on the stock-market valuation of the firm. American Economic Review pp. 436–461.

Borelli-Kjaer, M., Schack, L. M., Nielsson, U., 2021. #metoo: Sexual harassment and company value. Journal of Corporate Finance 67, 101875.

Charles, K. K., Guryan, J., 2008. Prejudice and wages: An empirical assessment of Becker's *The Economics of Discrimination*. Journal of Political Economy 116, 773–809.

Dahl, G. B., Knepper, M., 2022. Why is workplace sexual harassment underreported? the value of outside options amid the threat of retaliation, working paper.

Fee, C. E., Hadlock, C. J., Pierce, J. R., 2013. Managers with and without style: Evidence using exogenous variation. Review of Financial Studies 26, 567–601.

Friedman, M., 1962. Capitalism and Freedom. University of Chicago Press, Chicago.

Gande, A., Lewis, C. M., 2009. Shareholder-initiated class action lawsuits: Shareholder wealth effects and industry spillovers. Journal of Financial and Quantitative Analysis 44, 823–850.

Gantchev, N., Giannetti, M., Li, R., 2022. Does money talk? divestitures and corporate environmental and social policies. Review of Finance 26, 1469–1508.

Giannetti, M., Wang, T. Y., 2023. Public attention to gender equality and board gender diversity. Journal of Financial and Quantitative Analysis 58, 485–511.

Gormley, T. A., Matsa, D. A., 2011. Growing out of trouble? corporate responses to liability risk. Review of Financial Studies 24, 2781–2821.

Gorton, G. B., Grennan, J., Zentefis, A. K., 2022. Corporate culture. Annual Review of Financial Economics 14, 535–561.

Graham, J. R., Grennan, J., Harvey, C. R., Rajgopal, S., 2022. Corporate culture: Evidence from the field. Journal of Financial Economics 146, 552–593.

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

Guo, B., Pérez-Castrillo, D., Toldrà-Simats, A., 2019. Firms' innovation strategy under the shadow of analyst coverage. Journal of Financial Economics 131, 456–483.

Hart, O. D., 1983. The market mechanism as an incentive scheme. Bell Journal of Economics pp. 366–382.

Hayes, R. M., Jiang, F., Pan, Y., 2021. Voice of the customers: Local trust culture and consumer complaints to the CFPB. Journal of Accounting Research 59, 1077–1121.

Heese, J., Pérez-Cavazos, G., Peter, C. D., 2022. When the local newspaper leaves town: The effects of local newspaper closures on corporate misconduct. Journal of Financial Economics 145, 445–463.

Hersch, J., 1991. Equal employment opportunity law and firm profitability. Journal of Human Resources 26, 139–153.

Hersch, J., 2011. Compensating differentials for sexual harassment. American Economic Review 101, 630–634.

Hersch, J., 2018. Valuing the risk of workplace sexual harassment. Journal of Risk and Uncertainty 57, 111–131.

Kalev, A., Dobbin, F., 2006. Enforcement of civil rights law in private workplaces: The effects of compliance reviews and lawsuits over time. Law & Social Inquiry 31, 855–903.

Karpoff, J. M., Koester, A., Lee, D. S., Martin, G. S., 2017. Proxies and databases in financial misconduct research. Accounting Review 92, 129–163.

Kelly, B., Ljungqvist, A., 2012. Testing asymmetric-information asset pricing models. Review of Financial Studies 25, 1366–1413.

Kline, P., Rose, E. K., Walters, C. R., 2022. Systemic discrimination among large US employers. Quarterly Journal of Economics 137, 1963–2036.

Knight, C., Dobbin, F., Kalev, A., 2022. Under the radar: Visibility and the effects of discrimination lawsuits in small and large firms. American Sociological Review 87, 175–201.

Kumar, A., Page, J. K., Spalt, O. G., 2011. Religious beliefs, gambling attitudes, and financial market outcomes. Journal of Financial Economics 102, 671–708.

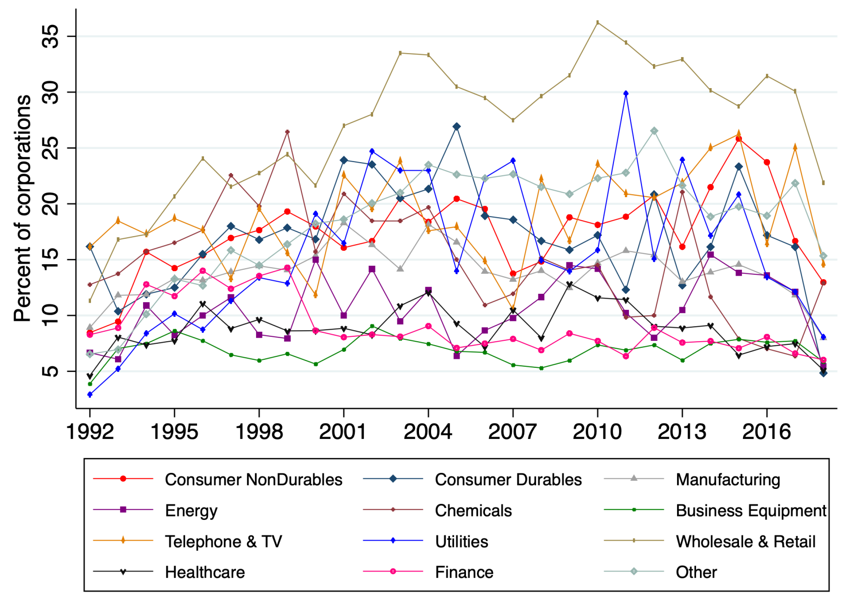
Lins, K. V., Servaes, H., Tamayo, A., 2017. Social capital, trust, and firm performance: The value of corporate social responsibility during the financial crisis. Journal of Finance 72, 1785–1824.

Liu, X., 2016. Corruption culture and corporate misconduct. Journal of Financial Economics 122, 307–327.

Parsons, C. A., Sulaeman, J., Titman, S., 2018. The geography of financial misconduct. Journal of Finance 73, 2087–2137.

Wang, T. Y., Winton, A., Yu, X., 2010. Corporate fraud and business conditions: Evidence from IPOs. Journal of Finance 65, 2255–2292.

Panel A. Industry distribution



Panel B. Geographic distribution

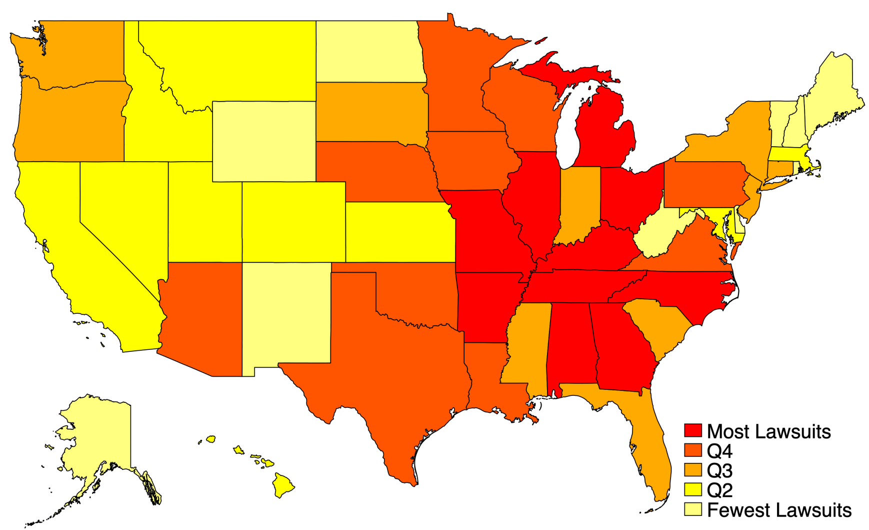


Fig. 1. **Distribution of discrimination lawsuits**
This figure displays the fraction of U.S public companies facing an employment discrimination lawsuit in federal court. The full sample consists of 94,113 firm-year observations from 11,331 U.S. public companies with data available in Compustat between 1992 and 2018. Panel A splits the sample by Fama-French 12 industry classification and calendar year. Panel B splits the sample by headquarter location.

Table 1: Characteristics of lawsuit filers and cases

This table reports descriptive statistics for civil-rights employment lawsuits filed against public corporations. The sample used to produce these figures is our Pacer sample. Columns correspond to the protected class alleged; "All" pools all filings. Continuous variables (age, wage, experience, number of plaintiffs) are shown as means and medians; wages are in U.S. dollars and ages/experience in years. Indicator variables (e.g., Male/Female, job title, nature of discrimination, reporting behavior, remedies) are shown as the percent of filers. Because cases may include multiple allegations, multiple protected classes, and multiple filers, percentages need not sum to 100. Group columns restrict the sample to filings by the corresponding protected-class (e.g., "Age" includes cases alleging age discrimination, "Origin" includes cases alleging national origin discrimination, etc.). FMLA denotes the Family and Medical Leave Act.

		Protected Class							
	All	Age	Disability	FMLA	Gender	Origin	Race	Religion	Other
Mean # Plaintiffs	1.03	1.03	1.01	1.02	1.02	1.10	1.05	1.08	1.06
Median # Plaintiffs	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Plaintiff Demographics									
Mean Age (Yrs)	48.04	52.20	47.14	46.71	45.10	48.95	45.62	46.32	47.03
Median Age (Yrs)	49.00	53.00	48.00	48.00	44.00	47.00	46.00	46.00	48.00
Mean Wage ($)	243,680	267,604	321,686	672,747	66,622	352,584	166,289	66,882	1,286,597
Median Wage ($)	42,000	50,000	40,000	36,000	39,020	40,000	36,540	30,450	44,840
Experience (Yrs)	8.66	13.57	9.07	7.61	6.84	7.51	8.26	6.04	8.03
Median Experience (Yrs)	5.00	11.00	6.00	5.00	4.00	5.00	5.00	4.00	5.00
Male (%)	40.13	47.62	42.89	31.02	18.41	51.71	51.02	53.95	45.81
Female (%)	60.64	53.62	57.26	69.48	82.29	48.97	50.27	47.18	55.73
Job Title									
Executive (%)	10.14	12.68	7.95	12.61	10.97	10.15	9.75	11.35	9.82
Manager (%)	25.67	31.58	21.07	25.22	26.88	25.36	26.47	21.75	25.64
Employee (%)	64.19	55.74	70.98	62.17	62.15	64.49	63.78	66.89	64.54
Nature of Discrimination									
Systemic (%)	41.82	45.09	35.66	37.59	51.28	47.32	50.90	39.98	39.21
Harassment (%)	47.08	45.87	33.01	32.73	64.65	62.80	59.15	57.63	47.16
Retaliation (%)	55.67	47.49	56.86	83.32	63.75	61.73	59.93	58.65	68.48
Hiring (%)	3.12	4.32	2.80	0.18	3.82	7.40	4.70	13.41	3.54
Pay/Benefits (%)	10.58	11.53	7.28	8.16	12.60	12.89	12.20	12.38	12.35
Promotion/Demotion (%)	13.75	18.17	7.52	6.35	17.20	25.29	23.58	25.79	11.69
Termination (%)	50.35	57.48	46.71	52.13	50.91	61.37	55.08	58.34	54.24
Other Details									
Still Employed (%)	4.72	4.33	3.18	2.30	5.21	4.26	6.44	3.43	3.35
Reported Before Filing (%)	78.98	75.02	78.17	79.72	82.40	77.70	79.96	76.18	78.19
Comparative Claims (%)	37.71	50.49	26.46	29.94	43.35	44.52	49.02	33.78	32.61
Reinstatement or Backpay (%)	56.45	61.92	57.59	70.47	58.12	58.10	57.85	55.02	54.58

Table 2: Incidence and concentration of discrimination lawsuits
This table summarizes how federal employment-discrimination lawsuits are distributed across U.S. public corporations between 1992 and 2018. The statistics show both average lawsuit incidence and the concentration of filings across firms sorted by the number of years sued, employee count, and asset size. The sample consists of 94,113 firm-year observations from 11,331 U.S. public companies with data available in Compustat.

Summary	
Average number of discrimination suits per firm-year	0.39
Percent of firms that face a discrimination suit	
Never	73.45%
One year	10.67%
Two to five years	10.11%
Six to ten years	3.14%
Eleven to fifteen years	1.30%
Sixteen to twenty years	0.72%
More than twenty years	0.61%
Percent of total discrimination suits filed against (by employee count quartile)	
Firms in the bottom 25th percent of employee count	0.43%
Firms in the 25th to 50th percent of employee count	2.00%
Firms in the 51st to 75th percent of employee count	9.59%
Firms in the top 25th percent of employee count	87.31%
Percent of total discrimination suits filed against (by assets quartile)	
Firms in the bottom 25th percent of assets	1.00%
Firms in the 25th to 50th percent of assets	5.10%
Firms in the 51st to 75th percent of assets	14.13%
Firms in the top 25th percent of assets	79.73%
Top 100 firms facing discrimination suits contribute to	
Percent of total discrimination suits	49.44%
Percent of total employment	23.94%
Percent of total assets	18.63%

Table 3: Year, industry, and area fixed effects in discrimination lawsuits
This table compares the explanatory power of time, industry, and geography for the incidence of federal employment-discrimination lawsuits. The dependent variable is an indicator equal to one if a firm faces at least one discrimination lawsuit in a given year. Each column reports a separate linear probability model with the indicated fixed effects: year, Fama–French 12 industry, economic area (EA), or combinations thereof. The table reports R^2 values and joint F-tests for the fixed effects in the single-dimension specifications, along with incremental F-tests for adding industry or area effects to the indicated baseline models. The sample contains 94,113 firm-year observations for 11,331 U.S. public firms from 1992 to 2018.

	(1) Year FE	(2) Industry (FF12) FE	(3) Area (EA) FE	(4) Year + Area FE	(5) Year + Industry FE	(6) Year + Ind + Area FE
Observations	94,113	94,113	94,113	94,113	94,113	94,113
R^2	0.0025	0.0313	0.0316	0.0342	0.0350	0.0594
Statistical tests:						
Year FE						
F-stat	8.949					
Critical value for $p < 0.01$	1.756					
Industry FE					**vs. (1)**	
F-stat		30.848			288.046	
Critical value for $p < 0.01$		2.249			2.248	
Area FE				**vs. (1)**		**vs. (5)**
F-stat			18.114	18.265		14.409
Critical value for $p < 0.01$			1.271	1.271		1.271

Table 4: Peer lawsuit rates and the incidence of lawsuits
This table relates a firm's lawsuit incidence to the lagged lawsuit rates of three peer groups: local non-industry peers in the same economic area, local same-industry peers in the same economic area, and non-local same-industry peers outside the firm's economic area. Columns 1–3 use the full FJC–Compustat panel and distinguish total, local, and non-local lawsuits. Columns 4–7 use the PACER subsample and distinguish economic, non-economic, executive, and employee lawsuits. Regressors $\text{Lawsuits}_{t-1}^{-i,a}$, $\text{Lawsuits}_{t-1}^{i,a,-j}$, and $\text{Lawsuits}_{t-1}^{i,-a}$ are scaled by 100, so coefficients correspond to a one-percentage-point change in the relevant peer lawsuit rate. All specifications are Poisson models with lagged firm controls and year fixed effects. t-statistics clustered by firm are reported in parentheses. The full sample contains 94,113 firm-year observations for 11,331 U.S. public firms from 1992 to 2018; the PACER sample contains 21,127 firm-year observations for 1,681 firms from 1999 to 2018.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Full sample			PACER sample			
	# Lawsuits	# Local Lawsuits	# Non-Local Lawsuits	# Econ Lawsuits	# Non-Econ Lawsuits	# Exec Lawsuits	# Employee Lawsuits
$\text{Lawsuits}_{t-1}^{-i,a}$	0.011***	0.023***	0.010***	0.011***	0.013***	0.008	0.014***
	(3.24)	(8.17)	(2.58)	(2.65)	(3.04)	(1.07)	(3.58)
$\text{Lawsuits}_{t-1}^{i,a,-j}$	0.002	0.005**	0.001	0.003*	0.002	0.001	0.002
	(1.00)	(2.47)	(0.75)	(1.87)	(1.60)	(0.38)	(1.32)
$\text{Lawsuits}_{t-1}^{i,-a}$	0.025***	0.004	0.028***	0.023***	0.025***	0.006	0.024***
	(4.44)	(0.61)	(4.56)	(3.77)	(4.30)	(0.92)	(3.82)
Observations	85,889	85,889	85,889	18,732	18,732	18,732	18,732
Pseudo R^2	0.535	0.216	0.542	0.285	0.333	0.133	0.322
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 5: Which firms face discrimination lawsuits?

This table reports estimates from Poisson regressions that relate the number of discrimination lawsuits to lagged firm, industry, and headquarter-area characteristics. The full sample consists of 94,113 firm-year observations and 11,331 U.S. public companies with data available in Compustat between 1992 and 2018. The PACER sample consists of 21,127 firm-year observations from 1,681 unique firms over the period 1999-2018 for which we have available PACER data. *t*-statistics clustered by firm are reported in parentheses. The symbols *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. Table A1 lists variable definitions.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Full Sample			PACER Sample			
	# Lawsuits	# Local Lawsuits	# Non-Local Lawsuits	# Econ Lawsuits	# Non-Econ Lawsuits	# Exec Lawsuits	# Employee Lawsuits
Log(Assets/Employee)	0.243***	0.345***	0.231***	0.173***	0.186***	0.222***	0.182***
	(5.74)	(8.75)	(4.91)	(3.19)	(3.64)	(3.75)	(3.54)
Log(Employees)	1.806***	1.302***	1.868***	1.596***	1.488***	1.246***	1.532***
	(26.02)	(25.46)	(24.07)	(15.74)	(15.43)	(14.42)	(15.64)
Age	0.022	0.036	0.017	-0.009	0.006	-0.122*	0.030
	(0.39)	(0.76)	(0.27)	(-0.12)	(0.07)	(-1.67)	(0.37)
Return	-0.011	-0.073**	-0.002	-0.001	-0.025	-0.065	-0.006
	(-0.58)	(-2.36)	(-0.12)	(-0.04)	(-1.12)	(-1.25)	(-0.23)
ROA	0.191**	0.228***	0.199*	0.008	0.108	-0.273**	-0.014
	(2.09)	(3.45)	(1.89)	(0.07)	(0.81)	(-2.20)	(-0.15)
Market-to-Book	0.125**	0.107***	0.121*	0.180**	0.149**	0.109	0.218***
	(2.27)	(2.63)	(1.93)	(2.56)	(2.03)	(1.45)	(3.64)
Sales Growth	-0.132***	-0.143***	-0.132***	-0.139**	-0.103**	-0.090	-0.138***
	(-4.36)	(-3.46)	(-3.89)	(-2.52)	(-2.43)	(-1.34)	(-2.89)
Leverage	0.097***	0.000	0.108***	0.077**	0.088***	0.043	0.084**
	(2.99)	(0.01)	(3.06)	(2.00)	(2.58)	(0.85)	(2.26)
Ind. Price-Cost Margin	-0.036	0.004	-0.039	-0.029	-0.067	0.054	-0.075
	(-0.60)	(0.08)	(-0.58)	(-0.42)	(-1.03)	(0.76)	(-1.01)
HQ Log(Income Per Capita)	0.020	-0.305***	0.037	0.006	0.014	0.006	-0.035
	(0.13)	(-3.21)	(0.23)	(0.04)	(0.09)	(0.05)	(-0.20)
HQ Unemployment Rate	0.002	-0.075	0.008	-0.002	-0.004	-0.002	-0.030
	(0.04)	(-1.26)	(0.12)	(-0.03)	(-0.07)	(-0.03)	(-0.42)
HQ % Black Pop.	0.048	0.117***	0.040	0.017	0.018	0.089	0.005
	(1.01)	(2.96)	(0.78)	(0.34)	(0.37)	(1.59)	(0.11)
HQ Rural Area	-0.097	-0.393**	-0.077	-0.089	-0.100	-0.452**	-0.080
	(-0.50)	(-2.24)	(-0.36)	(-0.42)	(-0.44)	(-2.22)	(-0.35)
HQ Racial Bias Index	0.138	0.175***	0.140	0.142	0.141	0.108*	0.110
	(1.62)	(3.78)	(1.48)	(1.47)	(1.48)	(1.70)	(1.07)
HQ Least Sexist State	-0.195	-1.327***	-0.120	-0.162	-0.259*	-0.118	-0.265
	(-1.38)	(-8.43)	(-0.80)	(-1.00)	(-1.69)	(-0.84)	(-1.56)

Observations	94113	94113	94113	21127	21127	20699	21104
Pseudo-R^2	0.54	0.23	0.54	0.29	0.34	0.17	0.33
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 6: Geographic concentration and local-environment exposure

This table asks whether the association between lawsuit incidence and headquarter-location proxies is stronger for firms whose operations are more geographically concentrated. The regressions follow the specifications in Tables 3 and 4, respectively, but interact the location-based proxies with an indicator equal to one if the firm has above-median geographic concentration, defined as subsidiaries in three or fewer states. A stronger interaction effect is consistent with headquarter location being more informative when a larger share of the firm's operations is concentrated near that location. All specifications include lagged firm controls and year fixed effects. t-statistics clustered by firm are reported in parentheses. The symbols *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Full Sample			PACER Sample			
	# Lawsuits	# Local Lawsuits	# Non-Local Lawsuits	# Econ Lawsuits	# Non-Econ Lawsuits	# Exec Lawsuits	# Non-Exec Lawsuits
HQ Most Racist State	-0.282	0.292	-0.351	-0.345	-0.440*	-0.273	-0.567**
	(-1.26)	(1.35)	(-1.41)	(-1.29)	(-1.78)	(-1.33)	(-2.24)
HQ Most Racist State × Concentrated Operations	0.572**	0.549**	0.569**	0.610*	0.757**	0.948***	0.926***
	(2.32)	(2.05)	(2.07)	(1.87)	(2.51)	(2.93)	(2.89)
Concentrated Operations	0.033	-0.088	0.036	-0.069	-0.059	-0.276***	-0.039
	(0.37)	(-1.08)	(0.37)	(-0.65)	(-0.57)	(-2.77)	(-0.39)
HQ Least Sexist State	-0.095	-1.495***	-0.035	-0.226	-0.230	-0.148	-0.256
	(-0.44)	(-5.09)	(-0.16)	(-1.00)	(-1.21)	(-0.84)	(-1.17)
HQ Least Sexist State × Concentrated Operations	-0.277	0.076	-0.282	0.193	-0.006	0.286	-0.082
	(-1.17)	(0.18)	(-1.13)	(0.63)	(-0.02)	(1.07)	(-0.30)
Observations	47,096	47,023	47,096	16,113	16,130	15,787	16,130
Pseudo-R^2	0.52	0.21	0.53	0.29	0.33	0.16	0.33
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 7: Variance decomposition of lawsuit outcomes across firm and CEO
This table reports AKM-style variance decompositions for seven lawsuit outcomes, measured as log(1 + dependent variable) at the firm-year level. For each dependent variable, we estimate linear models with firm and CEO fixed effects and decompose the total variance of the outcome into the shares associated with firm-year controls (Log(Assets/Employee), Log(Employees), Log(Age), Stock Return, ROA, Market-to-Book, Sales Growth, Leverage; all lagged one year), firm fixed effects, CEO fixed effects, and the residual. Entries are percentages of total variance. "Controls-only R^2" is the variance share explained by the observable firm characteristics in a model without fixed effects. The last panel summarizes the connectedness of the firm–CEO bipartite graph used to estimate CEO effects, reporting the estimation sample size, the size of the largest connected component in terms of observations, firms, and CEOs, the number of CEO movers (CEOs observed at ≥ 2 firms), and the share of the estimation sample in the main connected group.

	# Lawsuits	# Local Lawsuits	# Non-Local Lawsuits	# Econ Lawsuits	# Non-Econ Lawsuits	# Exec Lawsuits	# Employee Lawsuits
Controls-only model							
Controls-only R^2 (total variance)	22.93	3.88	21.62	11.38	18.14	3.69	15.40
Firm vs. CEO							
Controls	12.52	1.35	12.71	6.91	14.87	4.18	9.88
Firm FE	55.69	27.74	55.95	36.33	40.28	20.87	41.43
CEO FE	6.96	9.35	6.83	6.92	6.30	7.64	7.07
Residual	24.61	61.13	24.25	49.51	38.46	66.83	41.53
R^2 (sum of components)	75.17	38.44	75.49	50.16	61.45	32.69	58.38
R^2 (from model)	75.39	38.87	75.75	50.49	61.54	33.17	58.47
Connectedness of the Firm–CEO two-way FE graph							
Estimation sample size (Firm–CEO AKM)	35,698	35,698	35,698	15,551	15,551	15,551	15,551
Largest connected group (obs)	123	123	123	57	57	57	57
Firms in largest group	7	7	7	3	3	3	3
CEOs in largest group	20	20	20	9	9	9	9
Movers (CEOs with ≥ 2 firms)	40	40	40	17	17	17	17
Share of estimation sample in main group (%)	0.34	0.34	0.34	0.37	0.37	0.37	0.37

Table 8: External CEO Turnover, Transition Types, and Discrimination Lawsuits *Panel A* reports the number of external CEO turnover events classified by the discrimination status of the CEO's origin and destination firms. *Panel B* reports Poisson pseudo–maximum likelihood (PPML) difference-in-differences estimates of how external CEO turnover affects the number of discrimination lawsuits filed against a firm. Columns (1) and (2) report the average post-turnover effect across all external CEO turnover events. Columns (3) and (4) allow the post-turnover effect to vary by whether the incoming CEO moves from a high-discrimination firm to a low-discrimination firm or from a low-discrimination firm to a high-discrimination firm. All specifications include firm×cohort and year×cohort fixed effects, and standard errors are clustered by firm.

Panel A. Transition counts

	Destination firm	
Origin firm	Low discrimination	High discrimination
High discrimination	30	42
Low discrimination	38	57
Total classified external CEO moves	167	
Total external CEO moves	367	

Panel B. PPML difference-in-differences estimates

	(1) # Lawsuits	(2) # Lawsuits	(3) # Lawsuits	(4) # Lawsuits
Post CEO Turnover	0.147*	0.144*	-0.396**	-0.370*
	(0.083)	(0.085)	(0.150)	(0.151)
Post × High → Low			3.196***	3.079***
			(0.462)	(0.415)
Post × Low → High			-0.471**	-0.377*
			(0.176)	(0.177)
Observations	4422	4028	4422	4028
Pseudo-R^2	0.632	0.644	0.632	0.644
Firm-cohort FE	Yes	Yes	Yes	Yes
Year-cohort FE	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes

Table 10: Index reconstitution: Local-linear regression discontinuity around the Russell 1000/2000 cutoff

This table tests whether exogenous shifts in investor base induced by Russell index reconstitution affect subsequent discrimination lawsuits. Columns 1, 3, and 5 report first-stage estimates for Russell 2000 membership in July; columns 2, 4, and 6 report the corresponding 2SLS estimates for log(1 + #Lawsuits). The instrument is an indicator for falling on the Russell 2000 side of the May rank cutoff, τ. All specifications use triangular kernel weights, include separate linear trends in the running variable on either side of the cutoff, and include year fixed effects. The sample is limited to the pre-banding period, 1996–2006. t-statistics in parentheses are heteroskedasticity-robust and clustered by firm.

	(1) FS R2000 membership (July)	(2) 2SLS log(1 + # Lawsuits)	(3) FS R2000 membership (July)	(4) 2SLS log(1 + # Lawsuits)	(5) FS R2000 membership (July)	(6) 2SLS log(1 + # Lawsuits)
IV(R2000)		-0.099		-0.020		0.015
		(-0.30)		(-0.08)		(0.09)
τ	0.346***		0.456***		0.577***	
	(3.43)		(6.03)		(9.26)	
Rank	0.783***	0.076	0.826***	-0.028	0.691***	-0.131
	(4.75)	(0.20)	(6.56)	(-0.08)	(6.61)	(-0.50)
$\tau\times$ Rank	-0.783***	-0.036	-0.739***	0.047	-0.624***	0.104
	(-3.94)	(-0.09)	(-5.31)	(0.14)	(-5.50)	(0.44)
Observations	130	130	279	279	441	441
R^2	0.659	0.000	0.773	0.000	0.816	0.005
F-Statistic	11.80	11.80	36.38	36.38	85.71	85.71
Bandwidth	100	100	200	200	300	300
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Table 11: Brokerage closure exposure and lawsuits

This table tests whether quasi-exogenous reductions in analyst coverage affect subsequent discrimination lawsuits. The treatment variable is *Exposure*, defined as the firm-year share of distinct covering analysts whose brokerages close in year $t+1$. Column 1 verifies that exposure predicts a decline in future analyst coverage. Column 2 relates the same shock to future lawsuits over years $t+1$ to $t+3$. Column 3 reports a pre-trend specification using lawsuits in years $t-3$ to $t-1$. All specifications are Poisson regressions estimated on stacked event-time panels built from IBES analyst–firm–year coverage data, 1990–2012, and include firm-cohort, year-cohort, and closure-cohort fixed effects plus lagged firm controls. t-statistics in parentheses are heteroskedasticity-robust and clustered by firm.

	(1) # Analysts (t+1 to t+3)	(2) # Lawsuit (t+1 to t+3)	(3) # Lawsuit (t-3 to t-1)
Exposure	-0.071***	-0.041	-0.058
	(-4.91)	(-0.86)	(-1.01)
Observations	172,981	173,378	161,968
Clusters	2,810	2,813	2,717
Pseudo-R^2	0.721	0.735	0.736
Firm-cohort FE	Yes	Yes	Yes
Year-cohort FE	Yes	Yes	Yes
Closure-cohort FE	Yes	Yes	Yes
Firm × cohort FE	Yes	Yes	Yes

Table 12: Shareholder value implications of discrimination lawsuits

This table reports percentage and dollar cumulative abnormal returns (CARs) around the public filing of federal employment-discrimination lawsuits. Panel A reports mean and median CARs for the full sample of 55,235 lawsuits against 3,867 U.S. public companies with data available in Compustat and CRSP between 1992 and 2018. Panel B reports the same event-study estimates for subsamples defined by lawsuit outcomes, plaintiff type, defendant history, employee role, protected class, and the nature of the alleged misconduct. We test differences in means using t-tests and differences in medians using quantile regressions. The symbols *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively. Appendix A1 lists variable definitions.

Panel A: Full sample

	5-Day CARs (%)				**5-Day Dollar CARs ($M)**			
	Market Adj.		Fama–French 3-Factor		Market Adj.		Fama–French 3-Factor	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median
CAR	−0.041*	−0.039	−0.059**	−0.054	−25.710***	−1.137***	−23.305***	−1.137***
Observations	55,235		55,235		55,235		55,235	

Panel B: Sub-samples

By lawsuit outcome:			
(1) Dismissed	−0.016	−0.091***	17,494
(2) Settled	−0.057*	−0.101***	23,840
(3) Judgment	−0.038	−0.096**	13,901
Diff. (1–2)	0.042	0.010	41,334
Diff. (1–3)	0.023	0.005	31,395
Diff. (2–3)	−0.019	−0.005	37,741
By judgment result:			
(1) Plaintiff win	0.066	−0.106	887
(2) Defendant win	−0.063	−0.091*	10,242
(3) Unknown	0.022	−0.116	2,772
Diff. (1–2)	0.129	−0.015	11,129
Diff. (1–3)	0.442	0.010	3,659
Diff. (2–3)	−0.085	0.025	13,014
By plaintiff type:			
(1) Individual	−0.035	−0.094***	54,428
(2) EEOC	−0.319*	−0.328*	807
Diff. (1–2)	−0.284	−0.234	55,235

(table continues)

By defendant history:			
(1) 1st suit	−0.153	−0.243**	3,867
(2) 2–10	−0.122**	−0.161***	12,285
(3) 11+	−0.002	−0.073***	39,083
Diff. (1–2)	−0.031	−0.083	16,152
Diff. (1–3)	−0.151***	−0.170***	42,950
Diff. (2–3)	−0.120***	−0.088***	51,368
By employee role:			
(1) Exec	−0.037	−0.103	1,557
(2) Mgr	−0.086	−0.059	3,941
(3) Emp	0.024	−0.046	9,491
Diff. (1–2)	0.050	−0.044	5,498
Diff. (1–3)	−0.061	−0.057	11,048
Diff. (2–3)	−0.111	−0.013	13,432
By protected class:			
Age	−0.048	−0.077	4,923
Disability	0.062	−0.046	4,735
FMLA	−0.019	−0.057	949
Gender	0.008	−0.050	6,356
National Origin	0.153	0.058	2,054
Race	−0.022	−0.031	6,944
Religion	0.023	−0.081	914
Other	0.067	−0.034	1,058
By nature of discrimination:			
Systemic	−0.035	−0.055	7,794
Harassment	−0.026	−0.070	8,872
Retaliation	−0.009	−0.034	10,215
Hiring	−0.013	0.004	589
Pay/Benefits	0.122	−0.090	2,005
Promotion/Demotion	0.023	0.016	2,705
Termination	−0.035	−0.051	9,362
By plaintiff gender:			
(1) Male	0.032	0.03	5,194
(2) Female	−0.082*	−0.087**	8,069
(2) Diff. (1–2)	0.112	0.122**	13,151

Appendix A. Appendix

A.1. PACER Extraction Prompt

To construct case-level variables from unstructured legal text, we use GPT-4 with a prompt that asks the model to extract metadata from each employment-discrimination complaint filed in U.S. federal district court. The extracted fields include docket and filing information; observable plaintiff characteristics such as gender, race, age, job title, tenure, compensation, and employment status; the protected classes and forms of discrimination alleged; narrative summaries of the underlying conduct; the identities of accused individuals; internal reporting and grievance activity; indicators of comparative treatment, discriminatory language, and broader workplace patterns; and the remedies sought. When information is not stated in the complaint, the model records the field as "Not Mentioned," which helps distinguish true absence from missing disclosure.

> **Prompt:** You are an expert legal assistant trained in identifying key information from employment discrimination lawsuits filed in U.S. federal district courts. Your task is to read through the complaint and extract structured data useful for empirical research. Extract the following fields as completely and precisely as possible. Provide "Not Mentioned" if a field is absent.
>
> **Document Metadata**
>
> 1. Docket Number
> 2. Filing Date
> 3. Judge Name
>
> **Plaintiff Characteristics (if discernible)**
>
> 4. Plaintiff Gender
> 5. Plaintiff Race or Ethnicity
> 6. Plaintiff Age (or age range)
> 7. Job Title or Role at Employer
> 8. Employment Tenure (years/months)
> 9. Plaintiff's salary
> 10. Was Plaintiff still employed at time of filing? (Yes/No/Unclear)
>
> **Discrimination Allegation Details**
>
> 11. Protected Class Involved (List all that apply: Race, Gender, Age, Disability, Religion, Sexual Orientation, National Origin, Pregnancy, Other)
> 12. Nature of Alleged Discrimination (Check all that apply: Termination, Harassment, Promotion Denial, Unequal Pay, Retaliation, Failure to Accommodate, Other)

13. Summary of Alleged Conduct (2–4 sentence description; use complaint language where possible)
14. Names and Titles of Accused Individuals
15. Did Plaintiff report the behavior internally before filing? (Yes/No/Unclear)
16. Was the alleged discrimination ongoing or a single event?

Red Flag Indicators of Animus or Hostile Culture

17. Any direct discriminatory language or slurs quoted? (List terms if any)
18. Comparative claims included? (e.g., "similarly situated male employees were promoted")
19. Is there evidence of systemic patterns or multiple victims mentioned? (Yes/No/Unclear)
20. Does the complaint allege prior similar incidents at the firm? (Yes/No)

Remedies and Outcome Details (if available in complaint or docket)

21. Was reinstatement or back pay requested?
22. Any indication of prior internal grievance or investigation?
23. Desired Relief Stated (e.g., damages, promotion, policy change)

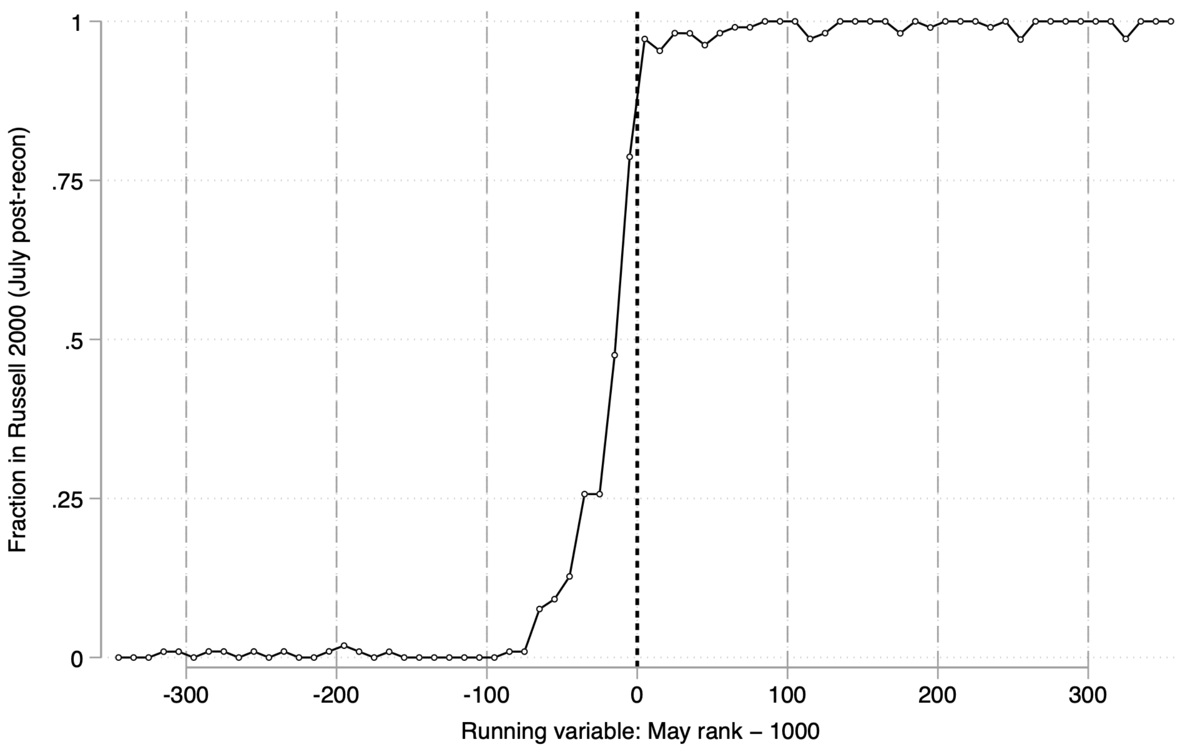


Fig. A1. **Russell 2000 membership around the index cutoff**
This figure plots July Russell 2000 membership after reconstitution against May rank relative to the Russell 1000/2000 cutoff, $\text{rank}_{\text{May}} - 1000$. Each marker shows the average July membership indicator in bins of width 10 ranks, using firm-year observations with May ranks in [650,1350] over the 1996–2006 reconstitution windows. The solid vertical line at zero marks the Russell 1000/2000 boundary, and dashed vertical lines indicate ± 100, ± 200, and ± 300 ranks around the cutoff. The sharp increase in membership probability at the cutoff illustrates the discontinuity exploited in the local-linear regression discontinuity design of Ben-David, Franzoni, and Moussawi (2019).

Table A1: Variable definitions.
This table lists variable definitions and data sources. CCM denotes the CRSP-Compustat Merged Database. CGP refers to Charles, Guryan, and Pan (2022). EXEC refers to the ExecuComp Annual Compensation Database. FJC denotes the Federal Judicial Center Civil Integrated Database. JR denotes John Romalis' website. KLD denotes the MSCI ESG KLD STATS Database. LLR refers to Levine, Levkov, and Rubinstein (2008). NBER refers to the NBER-CES Manufacturing Industry Database. WRDS denotes the WRDS Company Subsidiary Dataset.

Variable	Source	Description
# Lawsuits	FJC	Number of employment discrimination lawsuits filed in federal court against the firm during the year (Nature of Suit 442 – Civil Rights Job)
# Local Lawsuits	FJC	# of lawsuits where the plaintif filing county and firm HQ county are the same.
# Non-Local Lawsuits	FJC	# of lawsuits where the plaintif filing county and firm HQ county are not the same.
# Non-Econ Lawsuits	FJC	# of lawsuits where the lawsuit has low possibility of economic basis, which we define as lawsuits purporting harassment or retaliation against the plaintif.
# Econ Lawsuits	FJC	# of lawsuits where the lawsuit has a possible economic basis, which we define as lawsuits purporting no harassment or retaliation against the plaintif.
# Exec Lawsuits	FJC	# of lawsuits where the plaintiff is an Executive or Senior leadership position.
# Emp Lawsuits	FJC	# of lawsuits where the plaintiff is a rank and file employee.
Log(Employees)	CCM	Natural log of the number of employees (in thousands)
Log(Assets/Employee)	CCM	Natural log of the ratio of total assets to the number of employees
Age	CCM	Log number of years since firm was first listed in Compustat
Return	CCM/CRSP	Annual stock return; in tables this variable is labeled "Return"
ROA	CCM	Operating income before depreciation divided by total assets
Market-to-Book	CCM	Ratio of market value of assets to book value of assets
Sales Growth	CCM	Annual percentage change in sales
Leverage	CCM	Long-term debt plus debt in current liabilities divided by total assets
Ind. Price-Cost Margin	CCM	Industry-level price-cost margin, computed as sales minus cost of goods sold plus change in inventories, divided by sales plus change in inventories
HQ Log(Income Per Capita)	Census	County-level log income per capita for the firm's headquarter county
HQ Unemployment Rate	Census	County-level unemployment rate for the firm's headquarter county
HQ % Black Pop.	Census	County-level percentage of the population that is Black in the firm's headquarter county
HQ Rural Area	NCHS	Indicator equal to one if the firm's headquarter county is not designated as a Large Central Metro, Large Fringe Metro, or Medium Metro by the 2006 NCHS Urban Rural Classification Scheme, and zero otherwise
HQ Racial Bias Index	LLR	Difference between predicted and actual interracial marriage rates during 1970 in the firm's headquarter state
HQ Least Sexist State	CGP	Indicator equal to one if the firm's headquarter state is in the lowest two sexism ranking categories based on the General Social Survey, and zero otherwise

Appendix 1: Variable definitions (continued).

Variable	Source	Description
HQ Most Racist State	CGP	Indicator equal to one if the firm's headquarter state is in the top tercile of the HQ Racial Bias Index, and is zero otherwise.
Concentrated Operations	WRDS	Indicator equal to one if the firm has above median gegraphic concentration (subsidiaries in three or fewer states), and zero otherwise
$\text{Lawsuits}_{-i,a,t-1}$	FJC	Lagged lawsuit rate for local non-industry peer firms in the same economic area, scaled by 100 so coefficients correspond to a one–percentage-point change
$\text{Lawsuits}_{i,a,-j,t-1}$	FJC	Lagged lawsuit rate for local same-industry (excluding the firm) peer firms in the same economic area, scaled by 100 so coefficients correspond to a one–percentage-point change
$\text{Lawsuits}_{i,-a,t-1}$	FJC	Lagged lawsuit rate for non-local same-industry peer firms outside the firm's economic area, scaled by 100 so coefficients correspond to a one–percentage-point change
Post CEO Turnover	EXEC/CCM	Indicator that equals one in the year of a CEO turnover and all subsequent years, and zero otherwise
From High Discrimination	EXEC/FJC	Indicator that equals one when the incoming CEO previously worked at a firm with above-median average lawsuits per employee over the past five years, and zero otherwise
Post $\times$ From High Discrimination	EXEC/FJC	Interaction of Post CEO Turnover and From High Discrimination
R2000 Membership (July)	CRSP	Indicator equal to one if the firm belongs to the Russell 2000 index in July after the annual reconstitution (post–reconstitution membership), and zero otherwise
Rank	CRSP	May market-capitalization rank relative to the Russell 1000/2000 cutoff, constructed following Ben-David, Franzoni, and Moussawi (2019) and centered so that zero corresponds to the 1,000th largest stock by market capitalization; positive values correspond to stocks on the Russell 2000 side of the cutoff and non-positive values to stocks on the Russell 1000 side
τ	CRSP	Indicator equal to one if the firm's May market-capitalization rank is on the Russell 2000 side of the Russell 1000/2000 cutoff (i.e., $Rank > 0$), and zero otherwise
Exposure (Brokerage Closure)	IBES	Firm-year share of distinct IBES analysts covering the firm whose brokerages close in year $t+1$
Analysts	IBES	Number of distinct IBES analysts issuing earnings forecasts for the firm in a given year
5-Day CARs	CRSP	Percentage cumulative abnormal return over the 5-day window $[-2,+2]$ around a lawsuit filing, using either a market model or Fama–French three-factor model with a one-year estimation window ending 20 trading days before the event
5-Day Dollar CARs	CRSP	Dollar change in shareholder value over the 5-day window, computed as 5-Day CARs multiplied by the firm's equity market capitalization at the start of the event window

Table A2: Firm characteristics summary statistics

This table reports summary statistics (number of observations, mean, standard deviation, and selected percentiles) for firm-level, industry-level, and local demographic variables, as well as discrimination-lawsuit incidence measures, for the sample of 94,113 firm-year observations and 11,331 U.S. public companies with data available in Compustat between 1992 and 2018. Table A1 lists variable definitions.

	Mean	SD	P1	P25	P50	P75	P99
# Lawsuits	0.392	2.255	0.000	0.000	0.000	0.000	7.000
# Local Lawsuits	0.034	0.243	0.000	0.000	0.000	0.000	1.000
# Non-Local Lawsuits	0.340	2.096	0.000	0.000	0.000	0.000	7.000
# Econ Lawsuits	0.171	0.713	0.000	0.000	0.000	0.000	3.000
# Non-Econ Lawsuits	0.459	1.607	0.000	0.000	0.000	0.000	6.000
# Exec Lawsuits	0.053	0.279	0.000	0.000	0.000	0.000	1.000
# Emp Lawsuits	0.320	1.215	0.000	0.000	0.000	0.000	5.000
Log(Employees)	-0.076	2.049	-4.510	-1.599	-0.163	1.374	4.736
Log(Assets/Employee)	5.830	1.451	2.827	4.803	5.611	6.707	9.360
Age	2.566	0.827	0.693	1.946	2.565	3.219	4.111
Return	0.015	0.604	-0.958	-0.340	-0.069	0.215	2.961
ROA	0.060	0.183	-0.807	0.020	0.093	0.155	0.390
Market-to-Book	1.946	1.578	0.602	1.041	1.368	2.173	9.950
Sales Growth	0.127	0.359	-1.027	-0.016	0.085	0.225	1.767
Leverage	0.196	0.187	0.000	0.023	0.153	0.319	0.744
Ind. Price-Cost Margin	0.417	0.182	0.076	0.278	0.403	0.528	0.877
HQ Log(Income Per Capita)	10.608	0.408	9.842	10.306	10.583	10.857	11.785
HQ Unemployment Rate	0.300	0.238	-0.727	0.251	0.349	0.424	0.614
HQ % Black Pop.	0.138	0.116	0.004	0.048	0.106	0.200	0.558
HQ Rural Area	0.076	0.265	0.000	0.000	0.000	0.000	1.000
HQ Racial Bias Index	0.253	0.082	0.100	0.150	0.280	0.320	0.400
HQ Least Sexist State	0.105	0.306	0.000	0.000	0.000	0.000	1.000

Table A3: Firm characteristic differences This table reports mean values of firm-level, industry-level, and local demographic characteristics, as well as discrimination-lawsuit incidence measures, for five subsamples of U.S. public companies drawn from the 94,113 firm-year observations and 11,331 firms in our main sample (1992–2018). Columns report averages for the full sample, firms that ever face a discrimination lawsuit, firms that never face a lawsuit, firms with at least one PACER case, and firms with lawsuits but no PACER case. Table A1 lists variable definitions.

	Full Sample			PACER Sample	
	All	Has Lawsuits	No Lawsuits	Has Lawsuits	No Lawsuits
Log(Employees)	−0.076	1.302	−1.117	1.720	0.535
Log(Assets/Employee)	5.830	5.581	6.018	5.789	5.949
Age	2.566	2.846	2.354	2.996	2.880
Return	0.015	0.033	0.001	0.031	0.016
ROA	0.060	0.122	0.013	0.122	0.097
Market-to-Book	1.946	1.838	2.027	1.818	1.768
Sales Growth	0.127	0.107	0.142	0.081	0.082
Leverage	0.196	0.226	0.174	0.233	0.200
Ind. Price-Cost Margin	0.417	0.373	0.450	0.377	0.410
HQ Log(Income Per Capita)	10.608	10.619	10.599	10.776	10.778
HQ Unemployment Rate	0.300	0.282	0.313	0.275	0.263
HQ % Black Pop.	0.138	0.155	0.125	0.162	0.143
HQ Rural Area	0.076	0.060	0.088	0.061	0.055
HQ Racial Bias Index	0.253	0.265	0.244	0.266	0.256
HQ Least Sexist State	0.105	0.086	0.119	0.079	0.124
Observations	94,113	40,506	53,607	21,127	6,489