Non-Disclosure Agreements and Externalities from Silence*

Jason Sockin
University of Pennsylvania

Aaron Sojourner
University of Minnesota

Evan Starr
University of Maryland

January 18, 2022

Abstract

We examine how non-disclosure agreements (NDAs) influence firm reputation and the flow of labor market information by analyzing three 'NDA-narrowing' state laws that prohibited firms from using NDAs to silence workers regarding unlawful workplace conduct. We document three main results. First, these laws reduced average firm ratings by approximately 5% and increased the flow of negative information, as evidenced by workers providing more negative content in online employer reviews and a rise in work-related complaints to federal agencies. Second, these laws reduced the likelihood that workers conceal aspects of their identity when supplying negative information in online employer reviews—consistent with reduced concern about retaliation risks. Finally, we find suggestive evidence that ratings dispersion across firms increased within local labor markets, consistent with broad NDAs facilitating equilibria where firms with worse employment practices can 'pool' reputations among firms with better practices. Our results highlight how firms can use broad NDAs to preserve their reputations by silencing workers, but doing so imposes negative externalities on jobseekers who value such information and on competing employers who are less able to stand out.

JEL: M55, K31, J58
Keywords: Imperfect Information, Non-Disclosure Agreements, Firm Reputation, Externalities.

*We would like to thank David Hoffman, Alan Benson, Matthew Knepper, Kyle Rozema, and Ioana Marinescu for helpful comments and feedback. We would also like to thank seminar participants at the IZA Workshop on Labor Market Institutions, 4th IDSC of IZA/CAIS Workshop Matching Workers and Jobs Online, 14th Annual People and Organizations Conference, LERA 73rd Annual Meeting, and the University of Maryland. All mistakes are our own.

1 Introduction

A firm's reputation is important for competing in labor markets (Carmichael, 1984; Turban and Cable, 2003; Benson et al., 2020; Filippas et al., 2020; Sockin and Sojourner, 2020) and consumer markets (Shapiro, 1983; McDevitt, 2011; Cabral and Hortaçsu, 2010; Liu and Shankar, 2015). One way firms can manage their reputation is by prohibiting workers from disclosing negative information about the firm through the use of broad non-disclosure agreements (NDAs).[1] By silencing workers with negative information to share, NDAs may impose negative externalities on those who value such information, including job seekers, and competing employers attempting to differentiate themselves. These externalities have been spotlighted by the #MeToo movement; NDAs helped perpetuate violence by prohibiting victims of sexual harassment and assault from sharing information about the harms they experienced at work (Facchinei, 2020), reinvigorating the policy debate about whether such NDAs should be enforceable (Bast, 1999; Bagchi, 2015; Carlson, 2019).

No work to date, however, has empirically examined how NDAs influence firm reputation, the flow of information about employers, and whether NDAs give rise to these negative externalities. There are likely several reasons why. Data regarding the extent to which

[1] Sometimes firms use a 'non-disparagement agreement,' which prohibits workers from disparaging the firm. Since non-disclosure and non-disparagement restrictions are often commingled (Lobel, 2019; Drange, 2021), we use the term 'broad NDA' to refer to contracts that prohibit workers from sharing information generally. For examples, see Drange (2021) and `https://digitalcommons.law.scu.edu/cgi/viewcontent.cgi?article=2685&context=historical`.

In an article for the Society for Human Resource Management, employment attorney Grensing-Pophal (2019) advises firms to manage their reputation online by having all employees sign NDAs when hired. Then, if they share negative information online, they may be in violation. He says, "In my experience, a strong letter to the employee reminding them of their confidentiality obligations, sending them a copy of their agreement, and asserting the potential legal claims going forward may result in the employee removing the offending review."

Over 100 employers have sued Glassdoor to get the company to reveal the identity of reviewers. Though most suits failed, companies are clearly going after workers for posting unflattering reviews potentially in violation of their NDA (Glassdoor, 2018; Grothaus, 2020). In one example, Pariss Athena, a Black woman, discussed her experiences of racism while working at a large tech company on a podcast. "Almost immediately after the podcast was published, Athena asked the host to take it down. She'd forgotten she was under [an NDA] with the company she interned with and feared she could face legal retribution" (Birnbaum, 2020). Even Uber demands that its drivers sign NDAs to prevent the diffusion of bad information if they want the company's assistance dealing with negative experiences on the job, such as a carjacking (Kerr, 2021).

employers use NDAs are scarce, and changes to NDA enforcement regimes are rare.[2] In addition, because NDAs both censor (negative) information and protect against the leakage of trade secrets, it is difficult to separate two arguments for why NDAs might bolster firm reputation. One narrative is that NDAs actually make workplaces better, if firms profit from NDA-induced trade secret protection and share that value with workers. The other is that NDAs inflate firm reputation by discouraging workers from disclosing their bad, but not their good, experiences with an employer.

In this study, we leverage (i) novel data on NDA coverage, (ii) the recent adoption of state laws that prohibit firms from using NDAs to conceal unlawful activity—referred to herein as 'narrowing' NDAs—and (iii) employer reviews submitted by current and former employees on Glassdoor to examine how NDAs affect firm reputation and the flow of information workers volunteer about firms. Using double and triple-differences designs, our empirical approach examines how the relationship between NDA coverage and employer reviews change after states narrow NDAs relative to the same differences in states whose policies did not change.

Our findings strongly suggest that NDAs prop up firm reputation and suppress the provision of negative information. After the passage of NDA-narrowing laws, the overall ratings reported in employer reviews became roughly 5% more negative for an industry with average NDA usage, with stronger effects are observed among industries where NDAs are more prevalent. Additionally, the length of the 'cons' section of reviews increased 8% on average and descriptions related to harassment and workplace bullying rose 22%. These effects are driven mostly by first-time reviewers, are similar across several worker and firm characteristics, and resulted in increased ratings dispersion within firms. We also find evidence consistent with these three NDA-narrowing laws increasing workers' propensity to file complaints about unlawful workplace conduct with the Equal Employment Opportunity Commission (EEOC) and the Occupational Safety and Health Administration (OSHA).

[2] Hoffman and Lampmann (2019) note that, "We do not have a counterfactual firmly in hand. That is, to know what hush contracts do . . . , the gold standard test would be to find a legal regime that switched from enforcement to nonenforcement of hush contracts . . . . To date, such a natural experiment has been unavailable." (pp. 174–175)

Our interpretation of these results is that they reflect an uptick in the reporting of current or prior negative experiences rather than an increase in actual negative experiences per se. Narrowing NDAs likely discourages wrongdoing because bad behavior is now less shielded by NDAs. Indeed, we document short-run effects; over time, narrowing NDAs may improve employer quality if 'low-road' employers address the issues raised in negative reviews (Dube and Zhu, 2021), perhaps in response to subsequent hiring or retention difficulties.

While narrowing NDAs mitigates the negative externality on job seekers by reducing the censoring of negative information, the externality might persist if the value of such information to jobseekers is low. This can occur if workers who post negative reviews also conceal identifying job characteristics (e.g. their job title), which makes it difficult for jobseekers to ascertain the credibility or relevance of the supplied information. Indeed, Sockin and Sojourner (2020) show that reviewers who supply negative reviews are more likely to conceal aspects of their identity. Given this result, one might naturally expect that our finding of more negative reviews might lead to more identity concealment. However, the policies that narrowed NDAs also protected workers from potential retaliation for speaking out, raising the possibility of both more negative reviews and less identity concealment. Examining identify concealment, we find that narrowing NDAs reduces the likelihood of concealment by 20% on average. Consistent with our interpretation of narrowing NDAs reducing retaliation risk, these effects are stronger for reviews with more negative information or which mention harassment. Thus, policies that narrow NDAs attenuate the negative externalities on jobseekers by increasing both the supply and value of negative information.

With regards to potential mechanisms that underlie the rating and concealment results, we document the role of salience in the news media. While workers may be unaware of the laws themselves (Prescott and Starr, 2021; Rees-Jones and Rozema, 2019), we find that news coverage of NDAs peaks just before and during the passage of these laws, such that 20–30% of the effects we document for ratings and concealment can be attributed to salience. We also find similar negative but transitory effects from the Harvey Weinstein scandal, which predated passage of these laws by more than a year.

Finally, we also study whether policies that narrow NDAs allow workers to better distinguish between 'high-road' (i.e., worker friendly) and 'low-road' employers, where the reputation of low-road firms may be propped up by the suppressing nature of NDAs. To do so, we examine whether the variance of firms' average ratings changes after NDAs are narrowed for more vs. less NDA-intensive industries. While the results are somewhat noisy, we find that narrowing NDAs increases the standard deviation of reviews across firms within a market by 3% for an average NDA-incidence industry, and increases the inter-quartile range by 13%. This suggests that NDAs effectively compress the distribution of employer signals, preventing high-road employers from distinguishing themselves from low-road competitors.

Our results are, for the most part, robust to different measures of NDA exposure, different sets of control states, alternative weighting schemes, stacked designs, alternative ways of calculating standard errors, the potential for firms to farm positive reviews for themselves or plant negative reviews for their competitors, and to incorporating reviews where the location is unavailable. We also rule out that our results are driven by differences in wage growth.

Taken together, this work integrates across and contributes to several literatures in economics, law, and management. In the law literature, legal scholars spurred by extreme examples have been concerned about 'hushing contracts' because of the externalities they might impose on workers and firms, and frequently appeal to public policy arguments as a remedy (Bast, 1999; Bagchi, 2015; Hoffman and Lampmann, 2019; Note, 2006; Lobel, 2020; Yang and Liu, 2021).[3] However, despite these arguments and the media attention in the wake of the #MeToo Movement (Carlson, 2019; Griffith, 2021), only recently have scholars documented the ubiquity of NDAs (Starr et al., 2021; Balasubramanian et al., 2021), and no prior empirical work has examined the externalities they might induce. Thus, our core contribution is to provide the first causal evidence substantiating concerns about negative

[3]Note that the employer and employee who sign an NDA are unlikely to internalize the externalities fully because: (a) NDAs are often signed as a condition of employment before any information is realized, (b) all relevant parties, including future job seekers, are not involved in the contracting process, and (c) outside parties cannot price information of which they are not aware.

externalities that broad NDAs create by propping up firm reputation and suppressing information flows in labor markets. In doing so, we join other recent evidence that private contracting choices induce broad-based externalities (in the context of noncompetes: Starr et al. (2019); Johnson et al. (2020)). In turn, policies that enable workers to speak freely—and that limit the ability of firms to restrict disparaging speech generally (as opposed to only unlawful activity)—will further mitigate these negative externalities.[4]

Our results also contribute to the body of literature concerned with firm reputation (Diamond, 1989; Tadelis, 1999; Lange et al., 2011; Benson et al., 2020; Gadgil and Sockin, 2020), reputation management (Melo and Garrido-Morgado, 2012; Lii and Lee, 2012; Barrage et al., 2014; Lloyd-Smith and An, 2019), and labor market sorting (Sorkin, 2018; Sullivan and To, 2014; Maestas et al., 2018; Sockin, 2021). Respectively, these literatures highlight the importance of reputation as a meaningful asset for firms, the strategies firms use in order to maintain their reputations (e.g., advertising or corporate social responsibility), and the importance of non-wage amenities for labor market sorting. Our work is unique in highlighting how broad NDAs function as a reputation-preserving device for firms, likely increasing labor supply to such firms. However, NDAs do not just preserve reputations, but also compress the distribution of firm reviews, making it difficult for 'high-road' employers to stand out. Indeed, Sockin and Sojourner (2020) highlight that negative information about employers is both undersupplied and especially valued. Thus, our findings imply that prior work likely understates the importance of non-wage amenities in labor market sorting, since broad NDAs reduce the ability of workers to distinguish between firms.[5]

[4]In this sense, our work also relates to the literature studying worker voice (Hirschman, 1970), which typically considers either collective voice via unions (Mowbray et al., 2015) or workers expressing their dissatisfaction with management (Morrison, 2014). In the latter literature, a key challenge is gathering accurate information on why workers are discontent—our finding that current and low-tenure employees are less likely to speak out when they are bound by NDAs implies that NDAs may limit the ability of firms to discern what they might need to change. In addition, far less work in this literature has been concerned about voice 'to *anyone* who cares to listen' (Hirschman 1970 p. 4), e.g., potential job seekers, which is the main focus of this study.

[5]If information concealed by NDAs is revealed to potential hires via referral networks, however, then such workers may be largely exempted from the externalities precipitated by NDAs.

2 Institutional Background

NDAs are one in a class of employment provisions known as restrictive covenants, which seek to restrict what workers can do during and following the end of an employment relationship (Lobel, 2020). Other common restrictions include noncompetition agreements (Starr et al., 2021), which prohibit departing workers from joining or starting competitors, and nonsolicitation agreements, which prohibit departing workers from soliciting former clients or coworkers (Balasubramanian et al., 2021). These restrictive covenants can be agreed to at the outset of employment, while employed, or in severance arrangements or settlement agreements. NDAs typically prohibit the use or disclosure of trade secrets, but they can also "purport to protect information that is otherwise public, discoverable, or would not otherwise seem to be particularly confidential" (Flanagan and Gerstein, 2019). For example, in *Schwans Home Serv., Inc., 364 N.L.R.B. No. 20 2016* (page 6), the NDA stipulated that as a condition of employment:

> "Employee shall neither directly nor indirectly (i) disclose to any person not in the employ of Employer any Confidential or Proprietary Information, or (ii) *use any such information to the Employee's benefit, the benefit of any third party or [e]mployer, or to the detriment of Employer ...*" (emphasis added)

The NDA language in Weinstein Company, LLC employment contract provides another example.[6] It barred the employee from disclosing trade secrets and confidential information. The latter includes,

> "... any confidential, private, and/or non-public information obtained by Employee during Employee's employment with the Company concerning the personal, social, or business activities of the Company, the Co-Chairmen, or the executives, principals, officers, directors, agents, employees of, or contracting parties (including, but not limited to artists) with, the Company."

Language in non-disparage agreements, which are often written alongside or within non-disclosure agreements, directly prohibit the dislcosure of negative information about the

[6]The authors obtained this contract confidentially.

company. In data collected by Drange (2021), one non-disparagement agreement reads:

> "You promise that you will never directly or indirectly take any actions or make any statements, written or oral, that disparage or defame the goodwill or reputation of the Company, any affiliate or any of their directors, officers or employees."

Although legal scholars have long cautioned that NDAs may beget negative externalities (Bast, 1999), this work has been entirely theoretical, likely because data on NDAs were scarce until recently. In the first detailed analysis of NDAs—33,000 workers and 1,800 U.S. firms—Balasubramanian et al. (2021) find that NDAs are the most common restrictive covenant, covering 57% of the workforce and in use by 88% of firms for at least some workers. Moreover, they also show that NDAs are the baseline restrictive covenant: if a worker has (or the firm uses) a noncompetition or nonsolicitation agreement, there is at least a 95% chance that the firm also uses an NDA. Finally, they document substantial variation in NDA use. With regards to industries, construction has the lowest incidence among U.S. workers (41%) while professional, scientific, and technical services has the highest (70%).

Historically, among similar restrictive covenants, courts have been most willing to enforce NDAs (see Figure A1). This willingness derives both from the presumption that there is value for the two parties who have privately agreed not to share such information, and that NDAs impose comparatively weaker restrictions on workers relative to noncompetition and nonsolicitation agreements (Balasubramanian et al., 2021).

The NDA enforcement status quo remained largely stable until recently (Hoffman and Lampmann, 2019), when the #MeToo movement put a national spotlight on the externalities NDAs create. By silencing workers who experienced sexual harassment and assault, NDAs helped perpetuate such wrongdoing (Carlson, 2019; Silver-Greenberg and Kitroeff, 2020; Griffith, 2021).[7] The #MeToo movement—and the Catholic church sex abuse scandal before

[7] As Griffith (2021) writes about those seeking help from a firm that specializes in advocating for those speaking out about workplace problems, "almost all the firm's incoming clients had the same concern: Would they be sued for breaking their nondisclosure agreements? Such agreements were created by companies to protect valuable trade secrets, but they're also wielded as tools to keep employees from talking publicly about bad experiences at work."

that (Philip, 2002)—provide stark examples of negative externalities that arise in the context of NDAs: by preventing the diffusion of valuable, negative information, NDAs impose costs on both individuals not party to the agreement whom "bad actors" may harm in the future and on "good actors" who struggle to differentiate themselves from bad actors using NDAs.

As a result of the #MeToo movement and the essential role NDAs play in covering up and perpetuating wrongdoing, several states reconsidered how to regulate NDAs (Harris, 2019). While many states passed measures protecting workers from sexual harassment in the workplace or regarding post-harassment settlement agreements, only three states made it unlawful for firms to use NDAs (or other contracts) in the employment context to conceal any unlawful activity (Johnson et al., 2019). We detail these laws below.

Beginning on January 1, 2019, California SB 1300 made it an "unlawful employment practice, in exchange for a raise or bonus, or as a condition of employment or continued employment" to require an employee to sign "a nondisparagement agreement or other document that purports to deny the employee the right to disclose information about unlawful acts in the workplace, including, but not limited to, sexual harrassment."[8] The bill also prohibits retaliation against employees who do speak out in ways the new law allows. Importantly, the bill was designed to apply retroactively to all prior NDAs, and not just NDAs agreed to after the bill became law (Akopyan, 2019).

Similarly, Illinois Senate Bill 75, effective January 1, 2020, noted that "Any agreement, clause, covenant, or waiver that is a unilateral condition of employment or continued employment and has the purpose or effect of preventing an employee or prospective employee from making truthful statements or disclosures about alleged unlawful employment practices is against public policy, void to the extent it prevents such statements or disclosures, and severable from an otherwise valid and enforceable contract under this Act."[9] The bill also protects workers against retaliation, but it is not retroactive.

[8] For more details, see https://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=201720180SB1300.

[9] See https://www.ilga.gov/legislation/publicacts/101/101-0221.htm for the full bill.

Lastly, beginning on March 18, 2019, New Jersey Senate Bill 121 stipulated that a "provision in any employment contract" that would prohibit current or former employees from revealing "the details relating to a claim of discrimination, retaliation, or harassment" are "against public policy and unenforceable" (Hughes and Nacchio, 2019).[10] The bill also protects workers against any retaliation and makes the party attempting to enforce a contract against public policy liable for attorneys' fees. The bill does not apply retroactively.

Comparing each of these policies, the California policy is the broadest in terms of its applicability because it specifically highlights nondisparagement agreements (Legittino, 2019), covers all unlawful acts in the workplace, and, covers NDAs agreed to prior to the law's implementation (Akopyan, 2019). The Illinois law is less broad in that it only applies to 'unilateral' restrictions in new contracts after the effective date, carving out agreements for which there is some 'bargained-for-consideration.' The New Jersey law is the least broad because it only covers claims of discrimination, retaliation, or harassment, does not address other unlawful behavior, and is not retroactive. Although it is the least broad of the three, the New Jersey law is still much broader than all the other laws passed during this time because those focused almost exclusively on sexual harassment (Usenheimer et al., 2019).[11]

3 Data

3.1 Sources

To analyze how NDAs influence firm reputation and the information workers share about their employers, we use two unique datasets. The main dataset we leverage is from employer reviews submitted on Glassdoor from January 2015 through June 2021. This dataset is ideal for examining the effect of NDAs on the flow of labor market information, as it

[10] See `https://www.njleg.state.nj.us/2018/Bills/S0500/121_R2.PDF` for the full bill.

[11] For more details on these three laws and the weaker laws that passed, see Johnson et al. (2019). Given some of the heterogeneity of these laws, in our robustness checks we examine each state individually and also compare the results to those states that passed laws that covered predominantly sexual harassement.

consists of reviews written by those who would be bound to such NDAs—current and former employees—and because it is a common place prospective employees go to learn about employers. Workers are incentivized to leave reviews on the website through a "give-to-get" policy, whereby contributors to the website gain access to the corpus of information submitted by others. Reviews contain the worker's overall rating of their firm on a Likert scale of 1–5 stars, with more stars signaling more satisfaction. Each review also permits 1–5 stars ratings for five sub-categories (career opportunities, compensation and benefits, culture and values, senior management, and work-life balance), three (dis)approval options (CEO's performance, positive business outlook for the firm over the next six months, and would refer a friend to the firm), free-response text for 'pros' and 'cons' fields from which we can gain deeper insight into the content of the review, and an option to provide advice for management.[12] Volunteers are asked to provide their job titles and locations, but respondents can leave these identifying characteristics blank.[13] Summary statistics are provided in Table B1.

It is worth highlighting that while the Glassdoor data allow us to measure employer information flows in a systematic way, there is an important reason we should expect the relationship between NDAs and reviews on Glassdoor to be muted: because reviews on Glassdoor are anonymous they already offer workers some protection if they were to violate their NDA. However, this protection is not so strong as to suggest that we will not observe any relationship. Indeed, Glassdoor has been the subject of over 100 lawsuits in which firms have sought to reveal the identities of individuals posting negative comments (Glassdoor, 2018), and news stories highlight the potential for workers to be outed for their negative

[12]Appendix Figure A2 provides a sample, blank review form. Note that the language asks reviewers to "not post ... trade secrets/confidential information".

[13]The option to conceal one's job title was not always available to every potential reviewer. Figure A3 displays trends in job-title concealment rates and shows a structural decline in the rate after December 2018, reflecting changes to the Glassdoor review submission form on some user platforms. This coincides with the enactment of California S.B. 1300 and takes place a few months prior to the enactment of New Jersey S.B. 121. Fortunately, similar shifts are observed among both low- and high-NDA-use industries across both treatment and control states, minimizing concern that this could threaten identification. Note further that an exogenous reduction in the ability to conceal tends to cause an increase in average ratings supplied because it raises the expected retaliation risk of potential negative reviews (Sockin and Sojourner, 2020). In fact, we find the opposite. When the laws change, both concealment and ratings go down, as would follow from an exogenous reduction in legal and retaliation risk.

comments on Glassdoor (Grothaus, 2020). Indeed, lawyers regularly encourage firms to use broad NDAs as a way to discourage workers from leaving negative reviews specifically on Glassdoor (Grensing-Pophal, 2019).

We supplement the Glassdoor analyses with data on worker complaints made to two federal agencies, the Equal Employment Opportunity Commission (EEOC) and the Occupational Safety and Health Association (OSHA). These datasets allow us to measure how narrowing NDAs affects high-stakes claims about unlawful conduct at work.

The second dataset we leverage is from Payscale.com, from which we measure variation in the coverage of NDAs.[14] We aggregate the individual-level NDA data to calculate the average rate of NDA use by industry that we then merge with Glassdoor reviews. We use the industry level as opposed to the occupation or industry–occupation level, because, as we discuss later, volunteers may leave their occupation blank in the data as a form of identity concealment to protect against retaliation risk (Sockin and Sojourner, 2020), while industry follows from the always-observed firm. Nevertheless, alternative measures reveal similar results, which we document later in robustness checks. Harmonizing the Glassdoor and Payscale industry classifications results in fifteen industries. Table 1 lists these industries along with their respective shares of workers who report being bound by an NDA.

3.2 Cross-Sectional Analyses

This section provides descriptive evidence about how NDA incidence across industries relates to average employer ratings on Glassdoor (i.e., employer reputation). If NDAs suppress the amount of negative information shared across platforms, then we should see that average firm ratings are higher where NDAs are more common (because negative ratings are more likely to be missing).

[14] The data, developed in partnership with and discussed initially by Balasubramanian et al. (2021), derive from individual intake data collected by Payscale.com. In particular, individuals who visit the website can fill out information about themselves to receive an estimate of their earnings potential. In 2017, Payscale.com added a question on NDAs to their intake survey. Individuals were incentivized to provide accurate information because the validity of their earnings prediction depended on it.

Table 1: NDA Intensity by Industry in 2017

Industry	NDA incidence
Accommodation and Food Services	44.0%
Agriculture	51.0%
Arts and Entertainment	50.1%
Construction	41.3%
Finance and Insurance	70.1%
Health Care and Social Assistance	55.8%
Information	65.3%
Manufacturing	57.2%
Mining	59.6%
Other Services	48.3%
Professional, Scientific and Technical Services	70.4%
Real Estate	51.9%
Retail Trade	50.5%
Transportation and Warehousing	50.7%
Utilities	66.0%

Notes: The table provides the incidence of NDAs by industry per the unweighted, individual-level Payscale data, which cover 33,000 workers. See Balasubramanian et al. (2021) for more details.

For each industry, we calculate the average overall rating and relate these averages to each industry's share of workers covered by an NDA. There's a strong positive correlation (0.61) between the intensity with which NDAs are used and the average reported job satisfaction from workers in that industry.[15] To examine what variation is driving this relationship, we regress employees' overall ratings on the continuous measure of NDA incidence across industries, iteratively removing the variance across time, states, jobs, and workers. The results are displayed in Table 2 below. Column (1) indicates that jobs with a 10 percentage point higher likelihood of an NDA tend to have a 0.087 higher average employer rating on Glassdoor. This positive association moderates somewhat when controlling for state fixed effects (Column 2), and falls by approximately half when holding constant workers' job titles (Column 3). Finally, column (4) exploits the fact that individuals can leave multiple reviews for different employers with different likelihoods of using NDAs, allowing for the inclusion

[15] Figure A4 conveys this relation via a scatterplot.

of individual worker fixed effects. From this specification, we observe that the same worker gives more-positive reviews to firms in industries where they are more likely to be bound by an NDA.

Table 2: NDA Usage and Employees' Overall Satisfaction with Employers

	Overall rating			
	(1)	(2)	(3)	(4)
NDA intensity	0.874***	0.830**	0.457*	0.330***
	(0.293)	(0.281)	(0.240)	(0.094)
Year-month FE	✓	✓	✓	✓
State FE		✓	✓	✓
Job title FE			✓	✓
Worker FE				✓
Observations	3888393	3888393	3662024	352646
Adjusted R^2	0.02	0.02	0.06	0.40

Notes: The table examines whether there is a correlation between the intensity with which an industry use NDAs and the average job satisfaction of current and former employees within that industry. The dependent variable in each regression specification is the overall star rating (0-5) of the employer reviews. Standard errors are clustered by industry. Significance levels: * 10%, ** 5%, *** 1%.

As we noted above, the core challenge with these cross-sectional analyses is that NDAs may drive higher ratings for several reasons. They may prop up firm reputation by suppressing the flow of negative information, or raise actual employer quality through protecting trade secrets. Alternatively, it is also possible that the choice to use NDAs is correlated with many other characteristics that reflect firm quality (e.g., employing high-skilled, more-educated workers with high bargaining power). Accordingly, this analysis, while suggestive, can tell us little about whether NDAs actually improve firm quality, or whether the observed associations are a mirage driven by the suppression of negative information.

4 Empirical Approach

To address whether NDAs actually prop up firm reputation by causing workers to withhold negative reviews of their employers, we exploit the policy shocks in California, Illinois, and

New Jersey described above, which curtailed the ability of NDAs to conceal information about unlawful wrongdoing, including but not limited to discrimination, harassment, occupational safety, and retaliation. We refer to these policies as ‘narrowing NDAs,’ because they retain protection for trade secrets but narrow firms’ ability to suppress information about working conditions.

The passage of these three states’ policies makes a difference-in-differences design a natural fit, since many states did not pass these laws and industries differ in their NDA coverage. One intuitive specification is to look within high-NDA-use industries and compare the states that narrowed NDAs to those that did not. The key challenge with this approach, however, is that the bills that narrowed NDAs also came with a swath of other new laws such that there might be simultaneous treatments. These include practices geared towards addressing issues related to sexual harrassment in the workplace, such as limiting forced arbitration, prohibiting waivers of certain rights, and new training regimes (Johnson et al., 2019). Because these policies are designed to improve workplace quality generally (e.g., by reducing sexual harassment), a comparison of the states that narrowed NDAs to states that did not will be biased upwards because they will capture any increases in workplace quality due to these other policies and the effects of narrowing NDAs. Accordingly, to isolate the NDA-specific policies, we exploit industry-level heterogeneity in the likelihood that a worker is bound by an NDA to perform within-state analyses that net out any state-wide effects of these other, non-NDA-specific policies. Since the within-state analyses may be biased by industry-specific trends, in our preferred specification we leverage a triple-differences design to net out these differences using control states.[16] We revisit alternative specifications in Appendix B.

[16]Note that even if the other state policies have stronger effects in industries where NDAs are more prevalent, this would bias the effects of narrowing NDAs towards zero since the other policies are designed to improve workplace quality while we expect that narrowing NDAs will allow workers to reveal lower quality workplaces.

Accordingly, our main triple-differences specification is of the form:

$$Y_{ikst} = \beta \times NarrowedNDAs_{st} \times nda_{\iota(k)} + \lambda_{st} + \lambda_{\iota(k)t} + \lambda_{ks} + \epsilon_{ikst} \tag{1}$$

where Y_{ikst} is reported satisfaction for worker i employed at firm k in state s in calendar year–month t, $NarrowedNDAs_{st}$ is a dummy for whether one of the three NDA-narrowing laws was in effect in state s in year–month t, $nda_{\iota(k)}$ is the intensity with which NDAs are used in firm k's industry $\iota(k)$, λ_{st} are state-year-month fixed effects, $\lambda_{\iota(k)t}$ industry-year-month fixed effects, and λ_{ks} firm-state fixed effects. We two-way cluster the standard errors separately by state and industry, the two levels at which our key independent variables are assigned (Abadie et al., 2017)—though our results are robust to alternative methods for handling standard errors, including randomization inference.

In terms of control states, we would ideally include states whose trends reflect the counterfactual trends our treated states would have followed had they not narrowed NDAs. As this information is unobservable, we consider several alternatives. Since the pre-trends are fairly parallel in the model with all U.S. states included, we consider this our baseline model and later show that our results are not sensitive to the choice of control states.

In the triple-differences specification, our coefficient of interest β is identified by comparing how: (1) firm ratings change within the same industry-state around the enactment dates of laws narrowing NDAs, (2) in industries where NDAs are more- vs. less-prevalent, and (3) in states that passed such legislation compared with states that did not. If NDAs have a suppressing effect on firm ratings—by restricting the flow of negative content—then β should be negative, reflecting an increase in the flow of negative information after NDA-narrowing laws pass. The parameter $\frac{\beta}{100}$ describes the average effect of the legal change for a one percentage point increase in NDA intensity, such that β multiplied by the average level of NDA incidence (0.6) describes the average effect of prohibiting firms from using NDAs to conceal wrongdoing.[17]

[17] This design nets out any effects of other laws that were passed simultaneously which affect all industries

5 Results

5.1 Reputation and Information Flows

Our main analysis examines how ratings of employers—as a measure of firm reputation—change following the adoption of policies that narrowed NDAs. Table 3 builds to our main triple-differences specification by highlighting the baseline double-differences results. Columns (1) and (2) compare high-NDA-use industries in treated states to control states before and after NDAs are narrowed—where high-NDA usage refers to an industry with an above-average ($\geq$ 60.4%) incidence of NDAs. The results show that after NDAs are narrowed, the average firm rating among new reviews falls by 0.044 stars (or 1.2% of the sample mean) relative to control states. Columns (3) and (4) limit the sample to only the treated states of California, Illinois, and New Jersey, and examine how the relationship between NDA incidence and ratings changes after versus before NDAs are narrowed. The results suggest that narrowing NDAs reduces firm ratings by 3.8% of the sample mean (0.225*0.6/3.525) on average. This effect ticks up to 4.7% (0.270*0.6/3.478) when control states are incorporated into the model as a third difference (Columns 5–7). Overall, each specification tells the same story: the more likely workers are to be bound by an NDA, the more negative the average shift in ratings of their employers is following the passage of NDA-narrowing laws.

The key identifying assumption in these models is that the comparison group reflects the unobserved counterfactual of the treated group in the post-period. To assess whether the control group moved in parallel fashion in the pre-legislation period, Figure 1 reports the dynamic responses for the within-treated-state analyses (corresponding to Table 3: Column 4) and the specification with all states included (Table 3: Column 7), with the half-year before the laws are passed serving as the reference period. Before these policies take effect, we observe mostly parallel pre-trends, especially in the triple difference specification (panel

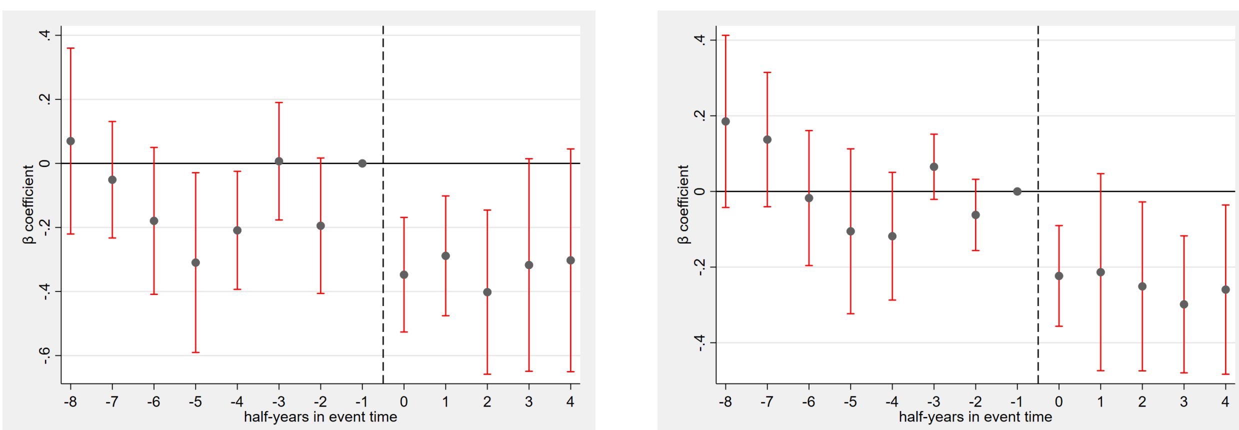
equally. The basic idea of this interpretation is that if NDAs cover 60% of the workers in an industry before NDAs are narrowed, then 0% of the NDAs can prohibit workers from sharing information related to unlawful conduct after NDAs are narrowed.

Table 3: Narrowing NDAs and Employees' Overall Ratings of Firms

	Full US Within High NDA Ind.		Within CA-IL-NJ High vs. Low NDA Ind.		Triple Difference		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
CA-IL-NJ	0.075*						
	(0.041)						
Narrowed NDAs	-0.034*	-0.044***			-0.014		
	(0.017)	(0.013)			(0.009)		
Narrowed NDAs x NDA intensity			-0.233*	-0.225*	-0.311***	-0.298***	-0.270***
			(0.132)	(0.122)	(0.001)	(0.050)	(0.079)
Dependent variable mean	3.583	3.583	3.525	3.525	3.488	3.488	3.478
Observations	1869291	1869291	818572	818572	3888391	3888391	3654296
Adjusted R^2	0.03	0.02	0.15	0.15	0.02	0.03	0.15
Year-Month FE	✓	✓	✓				
Industry FE	✓	✓					
State FE		✓					
Industry-Year-Month FE					✓	✓	✓
State-Industry FE					✓	✓	
State-Year-Month FE				✓		✓	✓
Firm-State FE			✓	✓			✓

Notes: The table implements difference–in–differences and triple-differences models estimating the causal effect narrowing NDAs had on the average overall rating of new employer reviews. High- and low-NDA use refers to industries for which the share of workers covered by NDAs is above- or below-average, respectively. Regressions in the first two columns are clustered by state, the next two clustered by industry, and the final three two-way clustered by industry and state. Significance levels: * 10%, ** 5%, *** 1%.

Figure 1: Narrowing NDAs and Employee Overall Ratings, Dynamic Responses



(a) CA-IL-NJ

(b) All states

Notes: The dependent variable in each regression is employee overall star rating. The sample period is 2015–2021 and point estimates are relative to the calendar half-year before the legislation goes into effect. Standard errors are clustered by industry in panel (a) and two-way clustered by industry and state in panel (b). Red vertical bars indicate 95% confidence intervals around each point estimate.

b), with declines in average ratings occurring in the post-legislation period.

While the analysis above examines the effect on average ratings, it is also illuminating to consider what effect narrowing NDAs has on the distribution of ratings. Glassdoor offers respondents five discrete options, with five stars representing the highest level of satisfaction and one star the lowest. We run linear probability models with indicators for each specific satisfaction level as the dependent variable, and examine where in the distribution of ratings the NDA-narrowing policies have the sharpest effect. The results are displayed in Table 4. Consistent with the idea that NDAs prevent workers from sharing negative information, we find that narrowing NDAs for an average industry decreases the likelihood of a five-star rating by 14.5% of the sample mean (0.078*0.6/0.323), while increasing the likelihood of a one-star rating by 16% (0.039*0.6/0.143), a two-star rating by 14% (0.026*0.6/0.113), and a three-star rating by 7% (0.021*0.6/0.189). Importantly, while the distribution clearly shifts downward after NDAs are narrowed, Table B2 shows that narrowing NDAs also increases the dispersion of ratings within firms—indicating hidden inequality in job satisfaction within firms that NDAs conceal.

Table 4: Narrowing NDAs and the Distribution of Review Ratings

	1(1 star)	1(2 stars)	1(3 stars)	1(4 stars)	1(5 stars)
Narrowed NDAs x NDA intensity	0.039***	0.026***	0.021**	-0.010	-0.078**
	(0.013)	(0.003)	(0.010)	(0.006)	(0.027)
Dependent variable mean	0.143	0.113	0.189	0.232	0.323
Observations	3654296	3654296	3654296	3654296	3654296
Adjusted R^2	0.09	0.02	0.04	0.03	0.16
Industry-Year-Month FE	✓	✓	✓	✓	✓
State-Year-Month FE	✓	✓	✓	✓	✓
Firm-State FE	✓	✓	✓	✓	✓

Notes: The table shows the triple-differences estimates relaying how the distribution of newly submitted employer reviews shifted between the five categorical options following the narrowing of NDAs. The dependent variable in each regression is a dummy for the specific star rating. Regressions include firm-state, industry-year-month, and state-year-month fixed effects. Standard errors are two-way clustered by industry and state. Significance levels: * 10%, ** 5%, *** 1%.

In Table B3, we show that the decline in reported sentiment is robust to considering alternative measures of worker satisfaction, and occurs in many domains. In columns (1)–(5) we examine the effects of narrowing NDAs on ratings in five sub-domains (e.g., senior management and career opportunities), and in columns (6)–(8), we consider three binary responses (e.g., whether the worker would refer a friend to the firm). For every measure, after new laws narrow NDAs, workers with a higher likelihood of being bound by an NDA report more negative information.

Reviews also contain free-response text for the volunteer to communicate the 'pros' and 'cons' of the workplace qualitatively to jobseekers. In Table 5, we examine several dependent variables reflecting the content of these fields. We find that the 'pros' share of the review's total text decreases by 2.4% (0.17*0.6/0.430) after NDAs are narrowed, driven by a 7.8% increase in the length of the 'cons' field. As a placebo test, we find no evidence that the pros section increased in length. Thus, individuals are not only giving firms lower ratings, they also spend a greater share of effort elaborating on the downsides of working at the firm.

Finally, because these laws were passed in the wake of the #MeToo movement, and all of the laws deal with harassment in some way, we examine whether individual reviews are more likely to use language related to harassment. Specifically, we create an indicator equal to one if any of the following terms are mentioned in the 'cons' field of the review: abus, assault, bully, bullied, harass, hostil, humiliat, innuendo, intimidat, mobb, sexual, stalk, threaten, victim, and violen. We then implement the same triple-differences estimation to see if reviews discuss harassment more frequently after these laws passed. Indeed, we observe that reviews indicating harassment increase by 22% (0.007*0.6/0.019) on average after NDAs are narrowed. In addition, workers do not just report more negative information, they are also 3% more willing to textually offer advice for management.

5.1.1 Extensive vs. Intensive Margin Results

The decline in average reviews could be due to a combination of several mechanisms operating along intensive and extensive margins. First, narrowing NDAs may have increased the number of reviewers or otherwise changed the composition of reviewers (on the extensive margin), so as to give firms more negative reviews on average—which would also increase

Table 5: Narrowing NDAs and Outcomes Related to Review Text

	Pros share of review text	Log length of Review text	Log length of Pros section	Log length of Cons section	Mentions harassment in review	Offers mgmt. advice
Narrowed NDAs x NDA intensity	-0.017***	0.043	-0.016	0.075***	0.007***	0.028***
	(0.004)	(0.041)	(0.058)	(0.019)	(0.002)	(0.008)
Dependent variable mean	0.430	5.297	4.431	4.553	0.019	0.562
Observations	3654296	3654296	3654296	3654296	3653412	3654296
Adjusted R^2	0.23	0.23	0.25	0.15	0.02	0.12
Industry-Year-Month FE	✓	✓	✓	✓	✓	✓
State-Year-Month FE	✓	✓	✓	✓	✓	✓
Firm-State FE	✓	✓	✓	✓	✓	✓

Notes: The table conveys how the content of worker reviews changed, according to the triple-differences specification, following the narrowing of NDAs. The dependent variable in each regression is listed as the header of each column. Regressions include firm-state, industry-year-month, and state-year-month fixed effects. For specification (1), each review is weighed by its character length. Standard errors are two-way clustered by industry and state. Significance levels: * 10%, ** 5%, *** 1%.

the within-firm dispersion of reviews, as we observe. It is also possible that individuals who would review regardless of the NDA laws were 'sweetening' their reviews, and now have the freedom to be more honest about their experiences with their employers. To examine these ideas, on the extensive margin, we look at changes in the volume of reviews, changes in the composition of reviewer characteristics, and the ratings of those reviewing on Glassdoor for the first time. On the intensive margin, we consider models examining within-reviewer changes in ratings across legal regimes.

To see whether the legal change increases the volume of reviews, we estimate our baseline difference-in-difference and triple-difference models using market review count as the outcome. All point estimates are consistent with an increase in the number of reviews, but estimates from almost all the models are imprecise as the analysis moves from review-level to market-level (Table B4).

The first six columns of Table 6 display the composition results. While reviewers are similar along many characteristics, after NDAs are narrowed reviewers are 5% (0.037*0.6/0.46) more likely to be female and 5% more likely to be long-tenured.

To assess potential extensive margin effects, in Column (7), we re-estimate our benchmark triple-differences specification for overall ratings but restrict the sample to only the first instances any particular worker submits a review on Glassdoor. The point estimate is very similar to our overall estimate. To examine intensive margin effects, in Column (8) of Table 6, we add individual volunteer fixed effects to our preferred triple-difference specification to consider whether the law affected ratings, holding the individual reviewer fixed (i.e., focusing on the subsample of reviews by volunteers who have left multiple ratings). Within worker, narrowing NDAs reduces average rating by 11% (-0.587*0.6/3.231). This result is consistent with these laws inducing workers to increasingly share negative information. Given that first-time reviewers comprise 91% of the sample, they drive most of the main results.

Table 6: Narrowing NDAs and Composition of Reviewers

	Indicator for worker type						Triple-difference	
	Current employee	Female	Short tenure	Ages 18–30	Post bachelors degree	Manager position	New reviewers	Include worker FE
Narrowed NDAs x NDA intensity	-0.020	0.037**	-0.049*	-0.032	-0.009	0.011	-0.252***	-0.587***
	(0.017)	(0.013)	(0.023)	(0.036)	(0.011)	(0.013)	(0.043)	(0.225)
Dependent variable mean	0.541	0.461	0.604	0.488	0.155	0.191	3.506	3.231
Observations	3654296	1766685	2891762	753748	653900	3152393	3275445	291880
Adjusted R^2	0.08	0.13	0.11	0.14	0.09	0.10	0.15	0.49
Industry-Year-Month FE	✓	✓	✓	✓	✓	✓	✓	✓
State-Year-Month FE	✓	✓	✓	✓	✓	✓	✓	✓
Firm-State FE	✓	✓	✓	✓	✓	✓	✓	✓

Notes: The dependent variable is an indicator variable for each worker worker type for Columns 1–6. In column 7, the dependent variable is the main rating (1-5) as in our benchmark triple-differences analysis, but incorporates worker fixed effects. Regressions are clustered by industry and state in Columns 1–6 and, due to convergence issues, by industry cross state in Column 7. Significance levels: * 10%, ** 5%, *** 1%.

5.1.2 Heterogeneity in Review Ratings

We also examine several dimensions of heterogeneity to explore the types of workers who are particularly likely to be silenced by NDAs. In terms of employee-level heterogeneity, we focus on the idea that narrowing NDAs is likely to have greater impacts on those who have more negative information to share and who experience the largest reductions in the probability of consequences for speaking out. This analysis leads us to focus on differences in whether the worker is a current or former worker and whether they are male or female. On the firm side, we focus on the idea that multi-state firms may be able to avoid the effects of these laws by stipulating other state's laws in their contracts (Coyle, 2020), and that it might be easier for smaller employers to identify reviewers who violate a broad NDA. Appendix C details these ideas further. While many of the estimates are in line with our theoretical expectations—the effects are more negative for current workers, firms that operate in one state, and smaller firms—the estimates are surprisingly consistent across each sub-sample, failing to pick up any statistically distinguishable effects.

5.1.3 Complaints to the EEOC and OSHA

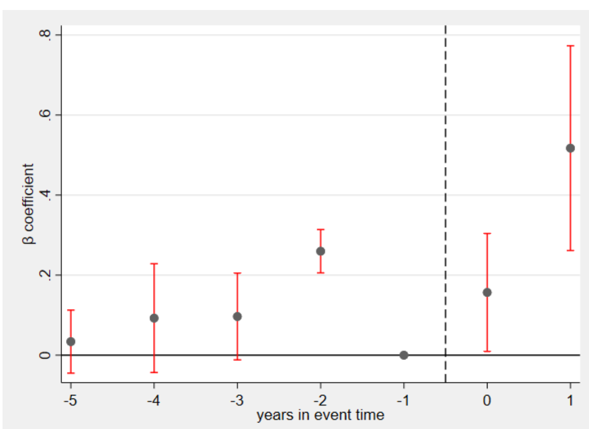
While the prior sections document how narrowing NDAs results in workers sharing more negative information about their employers on Glassdoor, in this section, we consider whether narrowing NDAs increases worker claims to federal agencies related to unlawful conduct in firms. We study the number of charges filed with the EEOC alleging sexual harassment and the number of complaint-driven, OSHA workplace inspections, which reflects the flow of credible complaints of occupational safety and health violations. The EEOC data are available at the gender-state-year (industry is not available). We aggregate the OSHA complaint-driven inspection data into counts by state-industry-year.[18] For the OSHA data, we estimate our benchmark triple-differences specification using a Poisson fixed effects specification, given that the outcomes are count variables. For the EEOC, we estimate a fully saturated Poisson model separately for men and women. Because the EEOC dataset is not further disaggregated by industry, this specification reflects only a double-differences model.

The dynamic event study plots for the EEOC and OSHA filings data are presented in Figures 2 and 3, respectively, while the overall estimates are presented in Table B5. Figure 2 shows that EEOC complaints rise for both men and women after NDAs are narrowed. On

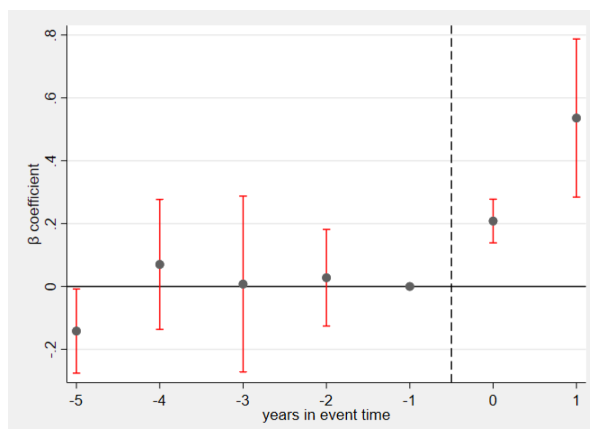
[18]The OSHA data can be downloaded from https://enforcedata.dol.gov/views/data_summary.php while the EEOC data is available at https://www.eeoc.gov/statistics/eeoc-fepa-charges-filed-alleging-sexual-harassment-state-gender-fy-1997-fy-2020. For OSHA, to focus on changes in worker reporting, we restrict attention only to complaint-driven inspections rather than inspections initiated for other reasons, such as in response to an accident or agency strategic initiative.

average, EEOC complaints rise 26% overall, driven by a 33% increase for male employees and a 16% increase for female employees. The differential for women is lower—and not statistically significantly different from zero overall—because sexual harassment cases for women rose the year before these laws were enacted, reflecting perhaps that these laws were a response to rising cases related to the #MeToo movement. Figure 3 similarly exhibits an uptick in OSHA complaints in the post-legislation period, although the fact the increase in complaints is lagged by a year renders the overall point estimate of 43% ($e^{(0.592*0.6)} - 1$) statistically indistinguishable from zero. The high point estimate perhaps reflects the fact that COVID-19 arose in the aftermath of the passage of these policies—giving workers more work-related issues over which they might file complaints. Overall, these results provide suggestive evidence that broad NDAs deter individual workers from raising claims of misconduct to official agencies.

Figure 2: Narrowing NDAs and EEOC Complaints, Dynamic Responses



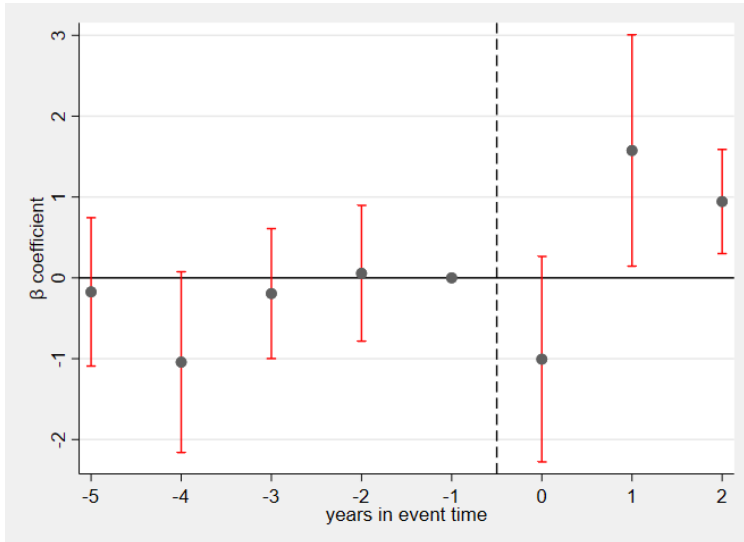
(a) Female filings



(b) Male filings

Notes: The dependent variable is the count of charges filed alleging sexual harassment by gender at the state-year level. The model was estimated via a fixed effects poisson model with fixed effects for state and year. The sample period is 2015–2020 and point estimates are relative to the calendar year before legislation goes into effect. Standard errors are two-way clustered by industry and state. Red vertical bars indicate 95% confidence intervals around each point estimate.

Figure 3: Narrowing NDAs and OSHA Filings, Dynamic Responses



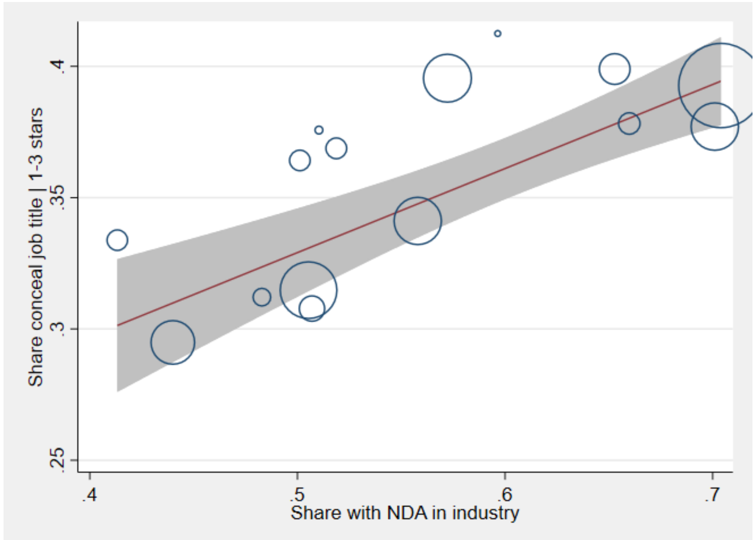
Notes: The dependent variable is the number of OSHA complaint-driven inspections at the state-industry-year level. The model was estimated via a fixed effects poisson model, with fixed effects for state-year, state-industry, and industry-year. The sample period is 2015–2021 and point estimates are relative to the calendar year before legislation goes into effect. Standard errors are two-way clustered by industry and state. Red vertical bars indicate 95% confidence intervals around each point estimate.

5.2 Job-Title Concealment

Up to this point, we have shown that laws narrowing NDAs make individuals more willing to give their firm low ratings and share more negative information. Since violating an NDA by sharing negative information invites potential legal costs, however, workers facing this risk may be especially likely to conceal their identity to mitigate the risk of employer retaliation. This identity concealment is important because consumers of a given review find it less valuable when volunteers conceal aspects of their identity, such as their job title or location (Sockin and Sojourner, 2020), perhaps because the information is less credible in general or it is more difficult to judge its relevance to one's own situation. As a result, even if narrowing NDAs makes workers more likely to share negative information, the value of that information would be attenuated if individuals supplying such reviews conceal their identity.

Retaliation protections built into the laws that narrowed NDAs may counteract the baseline incentive to conceal when sharing negative information, as it reduces the expected costs of any legal actions. If indeed narrowing NDAs reduces the likelihood of identity conceal-

Figure 4: NDA Usage and Rates of Job Title Concealment for 1–3 Stars Reviews



Notes: The figure illustrates the correlation between NDA incidence and average rate of job title concealment. Each dot represents an unique industry. Observations are weighted by the number of Glassdoor reviews within each industry. Industry-specific NDA intensity obtained from data through Payscale.com.

ment, then it would add more value to the additional (negative) information supplied.

To examine whether NDAs also increase identity concealment, we examine as a dependent variable whether the reviewer reveals their job title, conditional on giving a one-, two, or three-star review. Figure 4 examines the cross-sectional relationship between NDA intensity and the rate at which volunteers conceal their job title, conditional on the reviews containing a low rating. The figure shows a strong positive correlation (0.85) such that the likelihood of employees concealing their job title rises with the use of NDAs for negative ratings. This relationship may be driven, however, by other characteristics. For example, if industries in which NDAs are more likely to be deployed are generally more secretive, then reviewers in such industries may be less likely to leave such identifying information regardless.

To isolate the causal effect of NDAs on the likelihood of concealment, we bolster this cross-sectional analysis by examining how these patterns change when states narrowed NDAs using our main specification, but with a new dependent variable indicating job-title concealment. As in Table 3, in Table 7 we build up to our main specification for job title concealment, showing the baseline difference-in-differences results—leveraging variation within high-NDA

Table 7: Narrowing NDAs and Job Title Concealment Among Negative Reviews

	Full US Within High NDA Ind.		Within CA-IL-NJ High vs. Low NDA Ind.		Triple Difference		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
CA-IL-NJ	0.029***						
	(0.009)						
Narrowed NDAs	-0.017***	-0.020***			-0.017***		
	(0.006)	(0.005)			(0.002)		
Narrowed NDAs x NDA intensity			-0.282***	-0.282***	-0.033***	-0.049***	-0.060***
			(0.060)	(0.060)	(0.007)	(0.015)	(0.015)
Dependent variable mean	0.184	0.184	0.181	0.181	0.159	0.159	0.160
Observations	760189	760189	340626	340626	1717208	1717208	1569510
Adjusted R^2	0.06	0.07	0.14	0.14	0.07	0.07	0.13
Year-Month FE	✓	✓	✓				
Industry FE	✓	✓					
State FE		✓					
Industry-Year-Month FE					✓	✓	✓
State-Year-Month FE				✓		✓	✓
Firm-State FE			✓	✓			✓

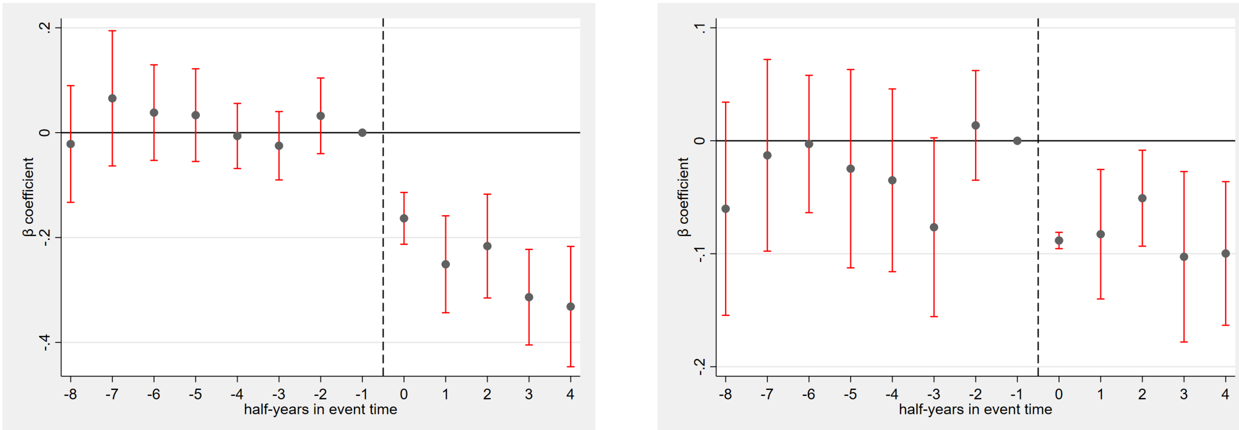
Notes: The table implements difference–in–differences and triple-difference models for estimating the causal effect narrowing NDAs on rate at which employees conceal their job title when leaving a negative review (of 1 to 3 stars). High- and low-NDA usage refers to industries for whom the share of workers covered by NDAs is above- or below-average, respectively. Regressions in first two columns are clustered by state, next two columns clustered by industry, and the final three columns two-way clustered by industry and state. Significance levels: * 10%, ** 5%, *** 1%.

industries over time and within treated states between industries with high and low NDA use—followed by our preferred triple-differences specification. Consistent with the cross-sectional results, after NDAs are narrowed, workers in industries where NDAs are more common become less likely to conceal their job title when supplying negative information relative to workers in industries where NDAs are less common (and relative to the same difference in states where NDA policies did not change). In our preferred specification (Column 7), narrowing NDAs decreases the likelihood of job-title concealment on average by 23% (0.060*0.6/0.160) relative to the sample mean.

To assess the key parallel trends identifying assumption in the concealment analysis, Figure 5 shows the within-treated-state analyses (corresponding to Table 7: Column 4) and the specification with all states (Table 7: Column 7). In both analyses, the pre-trends are relatively flat, followed by a drop in the post-period.

To develop additional evidence on the theorized mechanism (narrowing NDAs reduced

Figure 5: Narrowing NDAs and Job Title Concealment, Dynamic Responses



(a) CA-IL-NJ (Column 4)

(b) All states (Column 7)

Notes: The dependent variable in each regression is a dummy variable for the employee conceals their job title when submitting the review. Sample is restricted to negative reviews of 1–3 stars. The sample period is 2015–2021 and point estimates are relative to the calendar half-year before the legislation goes into effect. Standard errors are clustered by industry in panel (a) and two-way clustered by industry and state in panel (b). Red vertical bars indicate 95% confidence intervals around each point estimate.

retaliation risk which led to more-negative, less-concealed reviews being supplied), in Table 8 we examine whether concealment increases more in sub-samples consistent with that theory. Job-title concealment rates fall especially much for reviews where the 'cons' text is longer than the 'pros' text rather than vice-versa (for which we detect no statistically significant effect on concealment), and in reviews mentioning words associated with harassment in the 'cons' field rather than those that do not.

5.3 News Coverage as a Mechanism

Our results show that after California, Illinois, and New Jersey enacted legislation that narrowed NDAs, workers were increasingly more likely to rate their firms poorly, disclose more negative information, and be more less likely to conceal their identity. How exactly does this shift in policy lead to this newfound outpouring? There are several possible mechanisms. On the firm side, employers might change the actual content of their NDAs, or they may be less likely to enforce, either formally or informally, violations of NDAs which nevertheless are legal under the new law. On the worker side, news coverage pertaining to these new laws

Table 8: Narrowing NDAs and Job Title Concealment by Review Content

	Longer cons than pros	Longer pros than cons	Mentions harassment	Doesn't mention harassment
	(1)	(2)	(3)	(4)
Narrowed NDAs x NDA intensity	-0.060***	-0.037	-0.093**	-0.044**
	(0.008)	(0.023)	(0.040)	(0.015)
Dependent variable mean	0.145	0.133	0.194	0.138
Observations	1903765	1636487	42069	3579697
Adjusted R^2	0.12	0.13	0.14	0.12
Industry-Year-Month FE	✓	✓	✓	✓
State-Year-Month FE	✓	✓	✓	✓
Firm-State FE	✓	✓	✓	✓

Notes: The table implements difference–in–differences and triple-difference models for estimating the causal effect narrowing NDAs on rate at which employees conceal their job title when leaving a review with the specific characteristic noted in the column header. Standard errors are two-way clustered by state and industry. Significance levels: * 10%, ** 5%, *** 1%.

and NDAs more broadly might increase workers' awareness of their rights. Workers might also learn through their own social network or from coworkers that these laws have passed. Although we cannot observe firm-specific contracting policies, we can measure the extent to which NDAs are discussed in the news across states over time.

Following Rees-Jones and Rozema (2019), we obtain the number of articles that mention NDAs, as well as the respective word length of each article, in each state for each calendar month by scraping newslibrary.com.[19] If increased salience from elevated news coverage of NDAs is driving our results, then a necessary condition must first be observing an uptick in such news around these three laws. We estimate how NDA-related news evolves around passage of these laws by estimating the following dynamic difference-in-differences model,

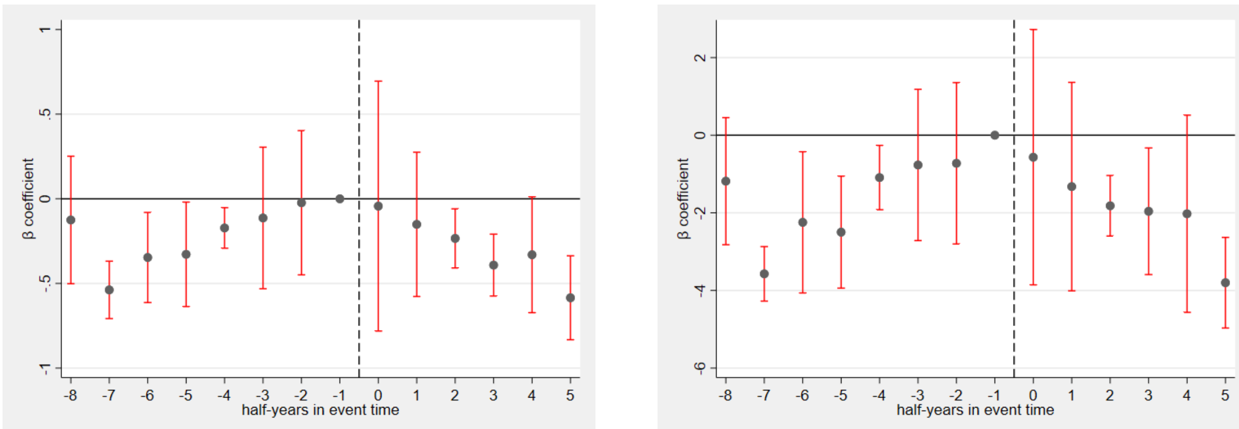
$$NewsNDAs_{st} = \beta_\tau \times NarrowedNDAs_{st} + \lambda_s + \lambda_t + \epsilon_{st} \tag{2}$$

where $NewsNDAs_{st}$ reflect the logarithm of total news articles or total words in said articles printed in state s in month t. Studying the time-varying coefficients β_τ presented in Figure 6 reveals that news coverage of non-disclosure agreements peaks for these three states in the

[19] In particular, we scraped articles including the term NDA, Non-disclosure, Nondisclosure, Non-disparagement, Nondisparagement, Confidentiality in the title.

year preceding the enactment of the law.[20] Relative to the six months prior to NDAs being narrowed, the number of NDA-related articles released two years before or after is 15–30% lower, corresponding to an 85—90% decline in words included in NDA-related articles.

Figure 6: Articles and Respective Word Counts Mentioning NDAs, Dynamic Responses



(a) Log number of articles

(b) Log number of words

Notes: The figures present dynamic difference–in–difference models estimating how the extent of news coverage pertaining to NDAs evolved after California, Illinois, and New Jersey passed laws narrowing NDAs compared with all other states. Data on the number of news articles discussing NDAs and the length of such articles are obtained at the state-month level following Rees-Jones and Rozema (2019). Red vertical bars indicate 95% confidence intervals around each point estimate.

Whether or not this uptick in NDA-related news contributes to our takeaway results can be directly tested by incorporating this variation in news exposure across states and time into our benchmark triple-differences model. Since news pertaining to NDAs peaks in the lead up to these laws being enacted, we consider lagged news coverage as potentially inducing more negative reviews. We then interact our lagged measure of NDA-based news coverage with our industry-level measure for NDA intensity (Table 1). The resultant triple-differences estimates on overall rating and the probability of concealing one's job title when providing a more-negative review are presented in Table 9. Accounting for the shift in news coverage attenuates the magnitude of our estimates—the overall rating estimate drops 20% (-0.270 to -0.215) while the job title concealment results falls 27% (-0.060 to -0.044) when two years of

[20]While the increased news coverage could correspond to discourse regarding the introduction of these bills in the state legislatures, or alternatively could reflect the zeitgeist of the period that motivated these laws, we abstract from clarifying news events as anything further than broadly referring to NDAs.

lagged news coverage is included. While this positive correlation between our policy shock and NDA-specific news is evidence supporting the causal interpretation of our treatment effect, it also highlights that heightened salience is not the only driving mechanism.

Table 9: Narrowing NDAs and Overall Ratings, Incorporating NDA News Coverage

	Overall rating			Conceals job title | 1–3 star reviews		
	(1)	(2)	(3)	(4)	(5)	(6)
Narrowed NDAs x NDA intensity	-0.270***	-0.244***	-0.215***	-0.060***	-0.048***	-0.044***
	(0.079)	(0.066)	(0.068)	(0.015)	(0.010)	(0.014)
Include one year lagged NDA news x NDA intensity		✓	✓		✓	✓
Include two years lagged NDA news x NDA intensity			✓			✓
Industry-Year-Month FE	✓	✓	✓	✓	✓	✓
State-Year-Month FE	✓	✓	✓	✓	✓	✓
Firm-State FE	✓	✓	✓	✓	✓	✓

Notes: The table implements the benchmark triple-differences model estimating the causal effect narrowing NDAs had on the average overall star rating and the likelihood of concealing one's job title when supplying a negative review when relative news coverage regarding NDAs across states over time is incorporated. Standard errors are two-way clustered by industry and state. Significance levels: * 10%, ** 5%, *** 1%.

5.4 Dispersion in Firm Reviews

Finally, we examine potential NDA-induced negative externalities borne by competing firms. In particular, because NDAs may inflate the reputation of low-road employers, it may be more difficult for high-road employers to credibly distinguish themselves in the labor market, in turn making it more difficult for job seekers to sort towards firms that differ in otherwise unobserved quality. While our heterogeneity analyses showed that narrowing NDAs caused workers to leave more negative reviews for firms with above-average ratings, this does not necessarily imply that better firms were able to differentiate themselves (i.e., if all highly reputable firms shifted down by the same amount). Accordingly, to examine this question, we consider whether dispersion in average firm ratings within a labor market rises after NDAs are narrowed. This would have implications for both firms and workers, as jobseekers may be more able to recognize firm qualities when the distribution of firm ratings becomes more dispersed. If, in contrast, the laws cause all firms to receive more negative reviews in equal measure, then dispersion would not change, nor would the ability of jobseekers to distinguish firms by relative quality.

We investigate this possibility by calculating the average rating for each firm k in state s in a given calendar half-year τ, $\bar{R}_{ks\tau}$. Then, given that firms map uniquely into an industry, we calculate the standard deviation across firms within each industry n in state s in each half-year τ, $\sigma_{ns\tau}$. This captures the degree of rating dispersion among the set of firms a worker might consider when searching for a job in a specific labor market (industry and state). We again estimate a triple-differences specification according to Equation 3,

$$\sigma_{ns\tau} = \beta \times NarrowedNDAs_{s\tau} \times nda_n + \lambda_{s\tau} + \lambda_{ns} + \lambda_{n\tau} + \epsilon_{ns\tau} \qquad (3)$$

where the outcome is the standard deviation of firms' average ratings for industry n in state s in calendar half-year τ. In our preferred specification, we weight each industry–state by the average number of firms represented in the market each half-year, such that dispersion in a market with five hundred firms is weighted ten times more than the dispersion in a market with fifty firms. Table 10 summarizes the results. Though the within-treated state results are imprecise, the results in the triple-differences setup suggest that dispersion in ratings across firms increases in high-NDA use industries following the passage of legislation narrowing NDAs. In our preferred specification (Column 8), narrowing NDAs increases the standard deviation of firm reviews by 3% (0.061*0.6/1.315) of the sample mean.

Figure 7 displays the dynamic effects according to the weighted within-treated-state analyses (corresponding to Table 10: Column 5) and the weighted specification with all states (Table 10: Column 8). In both cases, while there is some noise in the pre-period, dispersion in firms' average ratings rises in the post-period, especially in the triple-differences specification. Given the pre-period noise, we should interpret this result with due caution.

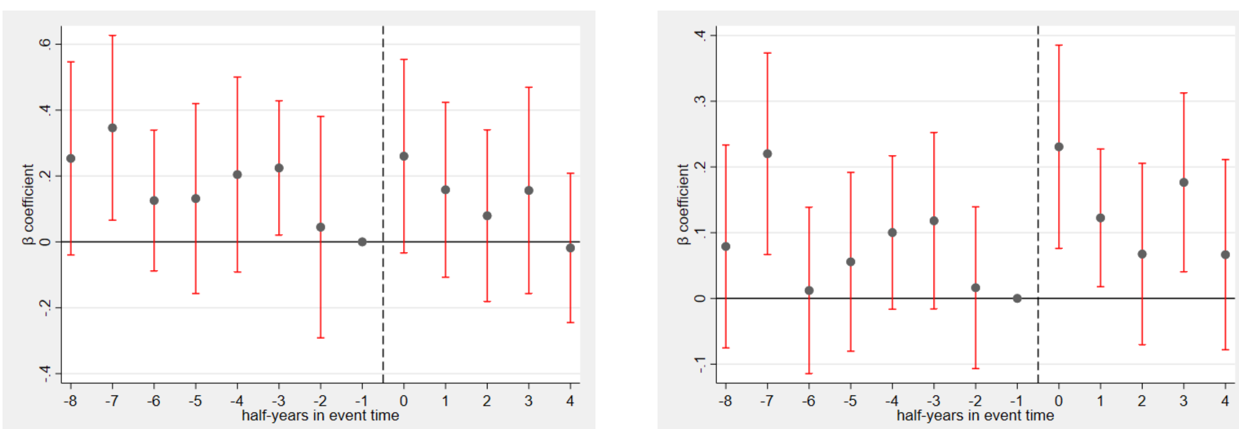
Given the noise in the pre-period and that the standard deviation of reviews may be highly skewed, we repeat this analysis using instead the inter-quartile range of average firm ratings in a labor market over time. That analysis, presented in Table B6, offers a similar

Table 10: Narrowing NDAs and Dispersion in Firms' Average Ratings

	Full US Within High NDA Ind.			Within CA-IL-NJ High vs. Low NDA Ind.		Triple Difference		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
CA-IL-NJ	-0.024***							
	(0.005)							
Narrowed NDAs	0.025*	0.025*	0.009*			0.011*		
	(0.014)	(0.013)	(0.005)			(0.006)		
Narrowed NDAs x NDA intensity				-0.027	0.133	0.118*	0.106*	0.061**
				(0.200)	(0.133)	(0.062)	(0.064)	(0.031)
Dependent variable mean	1.307	1.307	1.322	1.306	1.308	1.300	1.299	1.305
Observations	1534	1534	2648	468	581	4849	4823	9401
Adjusted R^2	0.39	0.48	0.53	0.36	0.53	0.66	0.67	0.62
Half-Year FE	✓	✓	✓	✓	✓			
State FE		✓	✓	✓	✓			
Industry-Half-Year FE						✓	✓	✓
State-Industry FE						✓	✓	✓
State-Half-Year FE							✓	✓
State-Industry 50+ firms on average	✓	✓		✓		✓	✓	
Weighted by average firm count			✓		✓			✓

Notes: The table implements difference–in–differences and triple-difference models for estimating the causal effect narrowing NDAs on the standard deviation of the (mean) firm rating within an industry–state pairing each calendar half year. High- and low-NDA usage refers to industries for whom the share of workers covered by NDAs is above- or below-average, respectively. Regressions in first three columns are clustered by state, next two columns clustered by industry, and the final three columns clustered by industry cross state. Significance levels: * 10%, ** 5%, *** 1%.

Figure 7: Narrowing NDAs and Dispersion of Firm Ratings, Dynamic Responses



(a) CA-IL-NJ, weighted (Col. 5)

(b) All states, weighted (Col. 8)

Notes: The dependent variable in each regression is the standard deviation of firms' average ratings within each industry–state–half-year. The sample period is 2015–2021 and point estimates are relative to the calendar half-year before the legislation goes into effect. Standard errors are clustered by industry in panel (a) and due to convergence issues, clustered by industry cross state in panel (b). Industry–state pairings are weighted by the average number of firms each calendar half-year. Red vertical bars indicate 95% confidence intervals around each point estimate.

conclusion. In our preferred specification, narrowing NDAs causes the interquartile range of firms' average ratings to increase 13% (0.469*0.6/2.078).

6 Sensitivity Analyses

In this section we consider several sensitivity analyses to probe the robustness of our results. Broadly speaking, we address the sensitivity of our results to alternative control groups, alternative measures of NDA intensity, alternative ways of handling standard errors, alternative methods for handling staggered adoption, alternative weighting schemes, the potential for firms to plant reviews, industry heterogeneity, state heterogeneity, and the issue of missing locations in reviews.

First, a recent literature highlights several concerns with staggered adoption in two-way fixed effects models (Goodman-Bacon, 2018; Callaway and Sant'Anna, 2020; Baker et al., 2021). Our results are unlikely to be affected by these issues because the policies we examine were all adopted within a year of each other and because we exploit within-state across-industry variation in the incidence of NDAs. Nevertheless, we employ a stacked regression approach to address the fact that some of our treated states serve as control states (Cengiz et al., 2019). To do this, we create a dataset with just one treated state and all of the control states. We then append another dataset with a different treated state and all of the control states. We repeat this for the third treated state, such that the data are stacked but that within each dataset, there is no variation coming from other treated states. We then implement our main triple-differences specification with fixed effects for each dataset. The result, presented in Column (1) of Table B7, reaffirms our main finding.

In Column (2) of Table B7, we address the fact that certain states and industries are more prevalent in the data. For instance, California represents 15% of the reviews sample, where as Illinois and New Jersey reflect 5.0% and 2.5% of the sample, respectively. To redistribute weight towards smaller states and industries, we consider an alternative specification, where industry-state pairings are given equal weight, meaning each review in industry n and state s has weight $1/\sum 1_{ns}$. Running our baseline triple-differences regression giving each industry–state equal weight does not change the takeaway result.

We also consider several alternatives to measuring NDA incidence. First, we construct alternative measures of NDA intensity from the Payscale data incorporating variation across occupations. Using the same survey from Payscale.com, we calculate the share of workers that are covered by NDAs within occupations and within industry–occupation pairs, where occupations reflect *onet50* occupation categories. To obtain occupations for Glassdoor reviews, we use a mapping from job title to occupation that was constructed based on Glassdoor's textual analysis machine learning algorithm. We then re-estimate a triple-differences model using a continuous measure of NDA intensity across industry-occupation pairs (Column 3) and occupations (Column 4). The results, though attenuated, remain robustly negative when utilizing these measures instead. Importantly though, these estimates include only those reviews with revealed job titles, which is an endogenous outcome itself. Therefore, we do not use either of these alternatives as our preferred specification.[21]

Second, we construct two firm-specific measures of NDA intensity again using data from Payscale.com. The first, in Column (5), leverages the individual-level data in Balasubramanian et al. (2021) and merges the employer-specific rates of NDA use with Glassdoor data by the name of the firm. The problem with this approach is that the Payscale dataset is not meant to be representative of firms and almost all firms in the sample have only one or two workers representing them. As a result, there is likely substantial measurement error. Even if the worker we sampled is not covered by an NDA, there may be another worker in the firm who is. Additionally, most firms in the Glassdoor data are not represented in the Payscale.com survey, an issue that is sidestepped by measuring NDA incidence at the industry level. Despite this challenge, our main result follows through under this alternative measure. The second, in Column (6), leverages an indicator for whether a firm uses an NDA

[21] This necessary sample restriction is non-trivial, as the sample sizes are cut by more than 40% when variation across occupations is incorporated.

with some or all of its workers, merging the firm-level data from Balasubramanian et al. (2021) into Glassdoor. The problem, however, is that only a tiny fraction of the firms that could be matched to Glassdoor reported using NDAs for none of their employees. Accordingly, the control group is particularly thin.[22] Nevertheless, our main result that the average overall rating falls among workers likely to be covered by NDAs after legislation that narrows NDAs is passed persists albeit with a somewhat smaller magnitude.

In Table B8, we consider how our main results change when we consider three alternative sets of control states. First, we consider regional neighbors only, under the assumption that neighbors may operate similarly to the treated states. Restricting our sample to only treatment states and these 'Neighbor' states, Column (2) shows that our point estimates from the rating, concealment, and dispersion analyses are unchanged. Second, we focus our attention on only states that are very well represented in the Glassdoor data. Restricting our sample to only treatment states and these 'High coverage' states, Column (3) highlights that our main results again hold under this comparison group. Finally, we consider the fact that several other states passed restrictions related to the #MeToo movement around this time period (see Johnson et al. (2019) for details). Restricting our sample to only treatment states and 'Weaker legislation' states, Column (4) reveals that our point estimates again are roughly unchanged, though the estimate for job title concealment does become statistically insignificant, perhaps due in part to the review sample thinning.

With regards to how we handle standard errors, Table B9 reports results from iterating over four different approaches for clustering standard errors, including only state, industry by state, firm and state (two-way clustered), and using the wild cluster bootstrap. Across all specifications the results hold. We next implement a randomized inference approach to gauge how often we could generate a result as negative as our triple-differences estimate from randomly allocating states between treatment and control groups. There are three possible treatments that can be assigned: January 2019, March 2019, and January 2020. We

[22] Recall that only 12% of firms do not use NDAs, per Balasubramanian et al. (2021), a number which falls even further when matching to Glassdoor.

randomly assign three states to these treatments pulling from a uniform distribution, assign the rest to the control group, and record the triple-differences estimate. After repeating this exercise five-hundred times, five percent of the simulation's coefficients fall below our true estimate of -0.270 (Figure A5).

Next, we consider the possibility that the decline we observe in overall ratings is driven by differences in wage growth in the post-legislation period. This issue poses a threat to the identification of our estimates because higher NDA intensity is strongly associated with higher pay—correlations of 0.62–0.69 across industries and occupations—and the three states that narrowed NDAs also increased their respective minimum wages around the same time.[23] If faster wage growth results in greater job satisfaction (Hamermesh, 1999), then our triple-differences estimates may reflect a wage effect rather than the narrowing of NDAs. To address this concern, we turn to earnings data available through Glassdoor.[24] Specifically, we calculate the average log earnings among full-time workers by calendar half-year within a given labor market, where a labor market reflects the pairing of a location (state/metro) and a job type (industry/occupation/industry–occupation), and incorporate this measure into our triple-differences model (Table B10). The robustly negative estimates from the narrowing of NDAs change little, reflecting that our observed effects appear unrelated to wage changes.

Our analysis thus far has been restricted to reviews for which a location is available, in order to assign reviews to treatment or control states. However, leaving the location of the review blank is not uncommon, representing nearly 41% of reviews. To attempt to incorporate these reviews into our analysis, we implement an imputation procedure by which reviews are assigned to their highest likelihood state. Although the location for these

[23] According to data from the Federal Reserve Economic Database (FRED), California raised the state minimum wage from \$11.00 in 2018 to \$12.00 in 2019, New Jersey raised the state minimum wage from \$8.60 in 2018 to \$10.00 in 2019, and Illinois raised the state minimum wage from \$8.25 in 2019 to \$10.00 in 2020.

[24] Other works have found Glassdoor pay data to be representative within industries and occupations. Karabarbounis and Pinto (2019) show the data broadly match first and second moments by industry and region using the Quarterly Census for Employment and Wages and the Panel Study of Income Dynamics, while Sockin and Sockin (2019) find correlations of about 0.9 and 0.8 for the first and second moments, respectively, by industry and three-digit occupation using the American Community Survey.

reviews is missing, the firm is not. The intuition behind our imputation process is to use all of the firms' reviews for which the location is not missing to estimate a latent distribution of the firms' reviews across states. We then assign every review for firm k with the location blank to the state s with the highest probability of origination, $p_{k,s} = \sum 1_{ks} / \sum 1_k$. If concealing location is not random and instead a strategic decision when revealing more-negative information, then failing to incorporate these reviews may bias our results. At the same time though, incorporating these no-location reviews injects measurement error that will likely bias our estimates toward zero. We re-estimate our baseline specification, incorporating reviews for which there is a reasonably-high probability the review is from state s, iteratively lowering the threshold $\bar{p}$ for inclusion into the sample, i.e. $p_{k,s} \geq \bar{p}$. Using a lower $\bar{p}$ introduces more measurement error into the state coding. While incorporating these reviews attenuates the magnitude of the effect on overall ratings, our estimates, displayed in Table B11, remain robustly negative.

We also consider the extent to which our results are driven by certain industries. We add to our main model of ratings and job title concealment industry fixed effects interacted with the post-legislation indicator. Figure A6 shows the results, where the y-axis conveys the post-legislation effect on overall ratings for each industry relative to zero and the x-axis the NDA incidence within that industry. Interestingly, industries with low-NDA incidence appear to have more positive ratings following the passage of this legislation. This may be due to the fact that California, Illinois, and New Jersey passed numerous other laws alongside the NDA provisions that sought to improve the workplace in light of the #MeToo movement, in addition to other policies such as raising the minimum wage. However, as in our main results, industries with high-NDA incidence, such as professional services and finance and insurance, have more negative reviews (and reduced likelihood of concealment). We also see a negative slope in both graphs, and it is this additional difference that nets out the effects of any non-NDA-interactive state policies that were implemented around the same time.

Next, we evaluate the extent to which our results are driven by either of the three state policies, which do differ from each other. We re-estimate our triple-differences specification, but isolating the effect for each policy separately by excluding the two other treated states entirely. The results, detailed in Table B12, reveal that California and Illinois have robust negative effects on overall ratings, while the results for New Jersey are statistically insignificant. These effects are not necessarily surprising, since California's policy was the broadest (and the only retroactive) policy and New Jersey's was the most narrow.

We also generate evidence against the possibility that our results can be attributed simply to a broader willingness to come forward with compromising information about low-road firms after the #MeToo movement. On October 5th, 2017, it was first reported by the New York Times that Harvey Weinstein had sexually harassed employees for decades, and this revelation ultimately led to accusations against a number of CEOs including those for Wynn Resorts, Guess?, and CBS.[25] If the #MeToo movement led to a large and persistent shift in workers' willingness to disclose negative information about firms that spread throughout industries prone to using NDAs, particularly in California, then our triple-differences estimate could reflect this change rather than the new legislation. There is a decline in overall rating among high-NDA industries in our treated states within three months following the revelation of the Weinstein scandal (Table B13). However, the effect fades. We find little evidence in support of a longer run effect that spills over into our treatment period.

Finally, we explore the potential role for firms strategically planting reviews as a way to bolster their own reputation relative to their competitors in explaining these results.[26] Could the observed empirical relationships be due to changes in insincere, mass review-planting induced by employers rather than decentralized behavior driven by workers? A

[25] See Weinstein, Wynn Resorts, Guess?, and CBS for the initial news coverage.

[26] According to Glassdoor's responses to frequently asked questions (FAQs) regarding content submission, the site takes a number of steps to prevent planted reviews. For one, Glassdoor requires an email address to make sure the respondent is "a real person" and respondents must "verify their account via email before any of their posts are shared." Further, Glassdoor makes a "commitment to review every post before it appears on the site." Nevertheless, some employers have been found to plant very positive reviews on the site (Fuller and Winkler, 2020).

firm that recently received a negative reputation shock due to the narrowing of NDAs might try to countervail the shock through increased self promotion or competitor degradation, or both (Mayzlin et al., 2014). If an employer copes with a NDA-narrowing induced negative shock to their reputation and plants more positive ratings of themselves, then the estimated effect of NDA-narrowing laws on average firm rating would be biased toward zero — working against finding a negative effect on average ratings. The fact that a negative effect is observed suggests that self-promotion is not driving our results. Alternatively, the negatively-shocked employer might have strengthened incentives to plant negative reviews for their competitors. In this case, our estimates would combine a direct negative shock to some firms' reputations driven by a change in workers' sincere reviewing behavior with those firms' strategic response to degrade its competitors.[27] In either case, both the self-promotion and competitor-degradation theories start by presuming that the employer is hit by a NDA-narrowing induced, worker-driven negative shock to reputation. Strategic self-promotion would dampen the effect while competitor degradation would amplify it.

To address these potential substitution patterns empirically, we use a proxy for detecting planted reviews following the methodology of Sockin and Sojourner (2020). The key idea is that employer-planted reviews will be more likely to occur as discontinuous spikes in the arrival rate of new reviews for a firm as employers engage in company-wide promotions or accumulative reviews prior to the announcement of awards recognizing employer quality on Glassdoor (Fuller and Winkler, 2020) — thereby breaking the prevailing trend. We consider various threshold growth rates in review volume to define suspicious spikes in review activity. Table B14 considers our main specification but investigates whether NDA-narrowing laws affect the probability of such spikes, at various growth rate thresholds for considering a

[27]Competitor-degrading is likely a more expensive, more difficult strategy than self-promotion. Each employer has multiple labor-market competitors, not all of whom are necessarily known to the employer. For an employer with C competitors to achieve the same gain in average review difference with their competitors, it would need to induce roughly C times as many negative reviews for competitors as positive reviews for themselves. Self-promotion is also almost certainly an easier task to accomplish. It requires only identifying one's own positively-disposed employees and encouraging them to post sincerely. In contrast, competitor-degradation requires either persuading employees to lie or finding competitors' negatively-disposed employees and encouraging them to post sincerely.

review to be suspect. The results suggest that these laws have little effect on this proxy for the number of planted reviews—–and that, if anything, the effect is negative. Importantly, our results are robust to the exclusion of reviews flagged for possible review planting (Table B15), evidence against our effects being fully driven by strategic review-planting behavior.

7 Conclusion

This study is motivated both by the longstanding concern and theory in the legal literature that firms can use overly-broad NDAs to prop-up their reputations by hiding negative information—resulting in negative externalities borne by workers and other firms (Bast, 1999; Hoffman and Lampmann, 2019)—and that prior research has not empirically examined whether such externalities arise. To this end, we leverage employee reviews on Glassdoor, variation in the incidence of NDAs, and three states' policy changes that prohibited firms from using NDAs to conceal unlawful activity or from retaliating against those who share such information. We document three primary results. First, broad NDAs prop up firm reputation and suppress negative information disclosure about the firm. Second, when employees do share negative reviews of their employer, NDAs make them more likely to conceal their identity. Prior evidence suggests this reduces the helpfulness of the information employees volunteer. Third, NDAs compress the distribution of firm reviews, making it harder for high-road employers to differentiate themselves from low-road employers whose reputations are propped up by NDAs. We now turn to our interpretation of these findings, along with their contributions, implications, and limitations.

A key limitation in interpreting these results is that we cannot observe when any wrongdoing occurred. That is, the fact that workers in high-NDA-use industries file more official complaints and provide more negative reviews of their employers after the passage of laws narrowing NDAs may suggest that these laws encourage more wrongdoing. While we cannot rule out this interpretation, we think it is unlikely for four reasons. First, negative information on employers is durable and can be shared at any point in time, even if the experiences occurred years ago. Second, the broadest law of the three we study is California's and it applies retroactively. Third, if firms know they cannot prohibit workers from speaking out by using NDAs, then they have less incentive to perpetrate wrongdoing in the first place. Fourth, our results are relatively short-run, giving firms limited time to react to these negative reviews. Indeed, firms do respond to negative Glassdoor reviews (Dube and Zhu, 2021), such that over the long run it is conceivable that firms address the issues raised in these negative reviews, resulting in the firms' ratings recovering.[28] Accordingly, our preferred interpretation of the results is that the policies narrowing NDAs encouraged workers to share their negative experiences—not that they created more negative experiences. We hope that as time elapses, future research examines the potential dynamic implications for firm behavior.

A second limitation is that, while we can show that salience in the news is responsible for 20–30% of the overall effects, we cannot pin down the mechanisms which explain the rest of the observed effect. Alternative mechanisms include that there are changes in the actual content of the NDAs workers are asked to sign, changes in firm enforcement protocol, or some other mechanism. In particular, recent research has shown that unenforceable contract terms can still chill worker behavior (Furth-Matzkin, 2018; Starr et al., 2020), in part because workers tend to believe that their contracts are enforceable, even when they are not (Prescott and Starr, 2021; Wilkinson-Ryan, 2017). In this context, broad NDAs may not be enforceable as written, but they may still silence workers who nevertheless are (i) concerned about breaking their contract, (ii) uncertain if the law covers the precise terms of their contract, or (iii) unaware of the laws that narrow NDAs. We hope that future work can sort out these various channels. Of particular import for future work is to gather a representative set of employment contracts to examine whether narrowing NDAs has any actual effect on the language written within contracts.

[28]As in Johnson (2020), revealed wrongdoing by one firm also may cause other firms to improve their behavior.

A third limitation is that—aside from our analysis of strategic review planting—we cannot examine the specific ways firms may seek to minimize negative information disclosures in the absence of protection from NDAs. Firms may, for example, hire more trusted employees, such as family members, or rely on other informal methods for preventing information disclosure—which may include alternative sticks or carrots. Our results nevertheless suggest that, despite any ways in which firms may strategically substitute to limit negative information flows, the net effect of narrowing NDAs still increases such flows.

A fourth limitation is that our data cannot distinguish precisely between non-disclosure agreements and non-disparagement agreements. Prior research suggests these are highly correlated (Drange, 2021), and broad NDAs certainly contain overlapping restrictions with non-disparagement agreements. Nevertheless, whereas non-disclosure clauses that protect trade secrets have some theoretical economic justification (Balasubramanian et al., 2021), this benefit is absent in the case of non-disparagement clauses, which simply prohibit employees from sharing negative information about the firm. Since the parties to non-disparagement clauses likely will not internalize the costs of silence on others, the negative externalities we identify here likely apply even more strongly to non-disparagement agreements.

A natural implication of our results relates to the literature on labor market sorting and reputation in labor markets (Carmichael, 1984; Chauvin and Guthrie, 1994). Prior research has found that non-wage amenities (including firm reputation, as in Bidwell et al. (2015)) account for a substantial portion of the value workers gain from a match (Sullivan and To, 2014; Sorkin, 2018; Maestas et al., 2018; Sockin, 2021). However, as we document, firms can use NDAs to conceal negative information about their jobs. As such, NDAs operate as a reputation-preserving mechanism firms can utilize—as long as laws allow it. Our findings suggest that NDA-narrowing laws increase negative information flows, which should impact labor market sorting. In many cases, firms with more negative ratings are less attractive to workers (Benson et al., 2020; Sockin and Sojourner, 2020).

In light of these ideas, we look for prima facie evidence of labor market effects on worker mobility and wages, using data from the Current Population Survey. Table B16 and Figure A7 show that we find small but noisy overall effects of narrowing NDAs on average wages and job transition rates. Lack of aggregate effects though does not necessarily mean that these policies do not have an effect. There are a number of competing mechanisms between information flow and wages and mobility. First, firms hit with a negative information shock may raise wages to compensate, while competitors might be able to lower them, resulting in a null effect overall. Second, NDA narrowing laws may also reduce any compensating differentials that were demanded for signing a broad NDA. Third, improved information for jobseekers might initially increase the probability of switching firms and might also improve the quality of any new matches, which in turn would reduce the probability of subsequent switches—creating countervailing effects on aggregate mobility.[29] Fourth, there may be an important time dimension: As matches improve over time from better information, productivity and wages may rise in the longer-run. However, this may require a longer post-treatment horizon to detect. Analysis of wages and mobility effects at the worker and firm level may prove fruitful in disentangling these mechanisms, but require different data. We leave this for future work.

Finally, while our focus in this study is on the externalities that NDAs create, policies that prohibit firms from using NDAs to conceal wrongdoing may come with important tradeoffs for the individuals who experienced the wrongdoing in the first place. Some individuals may not wish to share their negative experiences, and may prefer to receive a compensating differential in exchange for their silence. Given that these payments are typically private and endogenous to the expected or actual wrongdoing, they are necessarily difficult to study. Nevertheless, we hope that future research will engage not only with the potential externalities that NDAs create, but also with how much directly-harmed workers value their freedom to speak out—and how much they are compensated for giving up that freedom.

[29] Two recent court findings highlight that overly broad NDAs can effectively act as noncompetes (Tiku et al., 2021; Mertineit, 2020), suggesting that laws narrowing NDAs may spur mobility for workers bound by overly broad NDAs.

References

A. Abadie, S. Athey, G. W. Imbens, and J. Wooldridge. When should you adjust standard errors for clustering? Technical report, National Bureau of Economic Research, 2017. [Cited on page 15.]

H. Akopyan. California enacts legislation to combat discrimination and harassment. https://www.defenselitigationinsider.com/2019/02/06/california-enacts-legislation-to-combat-discrimination-and-harassment/, 2019. [Cited on pages 8 and 9.]

A. Bagchi. Other people's contracts. Yale J. on Reg., 32:211, 2015. [Cited on pages 1 and 4.]

A. Baker, D. F. Larcker, and C. C. Wang. How much should we trust staggered difference-in-differences estimates? Available at SSRN 3794018, 2021. [Cited on page 33.]

N. Balasubramanian, E. Starr, and S. Yamaguchi. Bundling postemployment restrictive covenants: When, why, and how it matters. Working paper, 2021. [Cited on pages 4, 6, 7, 11, 12, 34, 35, 42, and 56.]

L. Barrage, E. Chyn, and J. Hastings. Advertising and environmental stewardship: Evidence from the bp oil spill. Working Paper 19838, National Bureau of Economic Research, January 2014. [Cited on page 5.]

C. M. Bast. At what price silence: are confidentiality agreements enforceable. Wm. Mitchell L. Rev., 25:627, 1999. [Cited on pages 1, 4, 7, and 40.]

R. Beck. President biden?s proposed ban of (most) noncompetes: Protection strategies and steps to take now. https://www.faircompetitionlaw.com/2020/12/02/president-bidens-proposed-ban-of-most-noncompetes-protection-strategy-and-steps-to-take-now/, 2020. [Cited on page 49.]

A. Benson, A. Sojourner, and A. Umyarov. Can reputation discipline the gig economy?: Experimental evidence from an online labor market. Management Science, 66:1802–1825, 2020. [Cited on pages 1, 5, and 42.]

M. Bidwell, S. Won, R. Barbulescu, and E. Mollick. I used to work at goldman sachs! how firms benefit from organizational status in the market for human capital. Strategic Management Journal, 36(8):1164–1173, 2015. [Cited on page 42.]

E. Birnbaum. A wall of silence holding back racial progress in tech: Ndas. protocol, July 2020. [Cited on page 1.]

L. Cabral and A. Hortaçsu. The dynamics of seller reputation: Evidence from ebay. The Journal of Industrial Economics, 58(1):54–78, 2010. [Cited on page 1.]

B. Callaway and P. H. Sant'Anna. Difference-in-differences with multiple time periods. Journal of Econometrics, 2020. [Cited on page 33.]

G. Carlson. Gretchen carlson: Fox news, i want my voice back. https://www.nytimes.com/2019/12/12/opinion/gretchen-carlson-bombshell-movie.html, 2019. [Cited on pages 1, 4, and 7.]

H. L. Carmichael. Reputations in the labor market. The American Economic Review, 74(4): 713–725, 1984. [Cited on pages 1 and 42.]

D. Cengiz, A. Dube, A. Lindner, and B. Zipperer. The effect of minimum wages on low-wage jobs. The Quarterly Journal of Economics, 134(3):1405–1454, 2019. [Cited on page 33.]

K. W. Chauvin and J. P. Guthrie. Labor market reputation and the value of the firm. Managerial and Decision Economics, 15(6):543–552, 1994. [Cited on page 42.]

J. F. Coyle. A short history of the choice-of-law clause. U. Colo. L. Rev., 91:1147, 2020. [Cited on pages 22 and 63.]

D. W. Diamond. Reputation acquisition in debt markets. Journal of Political Economy, 97(4): 828–862, 1989. [Cited on page 5.]

M. Drange. We reviewed 36 ndas from major tech companies and discovered how far silicon valley's giants will go to silence and control their employees. Business Insider, 2021. [Cited on pages 1, 7, and 42.]

S. Dube and C. Zhu. The disciplinary effect of social media: Evidence from firms' responses to glassdoor reviews. Journal of Accounting Research, 2021. [Cited on pages 3 and 41.]

B. Facchinei. Gretchen carlson urges 2020 candidates to end ndas. https://www.newsy.com/stories/gretchen-carlson-it-can-be-suffocating-to-be-silenced/, 2020. [Cited on page 1.]

A. Filippas, J. J. Horton, and J. Golden. Reputation inflation. 2020. [Cited on page 1.]

J. Flanagan and T. Gerstein. 'sign on the dotted line': How coercive employment contracts are bringing back the lochner era and what we can do about it. USFL Rev., 54:441, 2019. [Cited on page 6.]

A. Fuller and R. Winkler. How companies secretly boost their glassdoor ratings. https://www.wsj.com/articles/companies-manipulate-glassdoor-by-inflating-rankings-and-pressuring-employees-11548171977, 2020. [Cited on pages 38 and 39.]

M. Furth-Matzkin. The harmful effects of unenforceable contract terms: Experimental evidence. Ala. L. Rev., 70:1031, 2018. [Cited on page 41.]

S. Gadgil and J. Sockin. Caught in the act: How corporate scandals hurt employees. (3774639), June 2020. [Cited on page 5.]

Glassdoor. Glassdoor free speech legal fact sheet: U.s. courts have almost always ruled that glassdoor reviews are protected opinion. Technical report, 2018. [Cited on pages 1 and 10.]

A. Goodman-Bacon. Difference-in-differences with variation in treatment timing. Technical report, National Bureau of Economic Research, 2018. [Cited on page 33.]

L. Grensing-Pophal. How to protect your employer brand on glassdoor. Society for Human Resource Management, October 2019. [Cited on pages 1 and 11.]

E. Griffith. How the world learns about bosses behaving badly. https://www.nytimes.com/2021/06/05/business/nondisclosure-agreements-lioness.html, 2021. [Cited on pages 4 and 7.]

M. Grothaus. Every employee's worst nightmare, getting outed on glassdoor, could become a reality. Wired, February 2020. [Cited on pages 1 and 11.]

D. S. Hamermesh. The changing distribution of job satisfaction. Working Paper 7332, National Bureau of Economic Research, September 1999. [Cited on page 36.]

E. Harris. Despite #metoo glare, efforts to ban secret settlements stop short. https://www.nytimes.com/2019/06/14/arts/metoo-movement-nda.html, 2019. [Cited on page 8.]

A. O. Hirschman. Exit, voice, and loyalty: Responses to decline in firms, organizations, and states, volume 25. Harvard university press, 1970. [Cited on page 5.]

D. A. Hoffman and E. Lampmann. Hushing contracts. Wash. UL Rev., 97:165, 2019. [Cited on pages 2, 4, 7, and 40.]

P. Hughes and M. Nacchio. New jersey bars nondisclosure agreements in all discrimination settlements, purports to bar arbitration agreements. https://www.natlawreview.com/article/new-jersey-bars-nondisclosure-agreements-all-discrimination-settlements-purports-to, 2019. [Cited on page 9.]

A. Johnson, K. Menefee, and R. Sekaran. Progress in Advancing Me Too Workplace Reforms in 20statesby2020. National Women's Law Center, 2019. [Cited on pages 8, 9, 14, 35, and 57.]

M. S. Johnson. Regulation by shaming: Deterrence effects of publicizing violations of workplace safety and health laws. American economic review, 110(6):1866–1904, 2020. [Cited on page 41.]

M. S. Johnson, K. Lavetti, and M. Lipsitz. The labor market effects of legal restrictions on worker mobility. Available at SSRN 3455381, 2020. [Cited on page 5.]

M. Karabarbounis and S. Pinto. What can we learn from online wage postings? evidence from glassdoor. Economic Quarterly, 104:173–189, 01 2019. [Cited on page 36.]

D. Kerr. Uber requires nondisclosure agreement before helping carjackeddriver. The Markup, August 2021. [Cited on page 1.]

D. Lange, P. M. Lee, and Y. Dai. Organizational reputation: A review. Journal of management, 37(1):153–184, 2011. [Cited on page 5.]

T. Legittino. California employment law changes you need to know. https://fkks.com/news/california-employment-law-changes-you-need-to-know, 2019. [Cited on page 9.]

Y.-s. Lii and M. Lee. Doing right leads to doing well: When the type of csr and reputation interact to affect consumer evaluations of the firm. Journal of Business Ethics, 105:69–81, 06 2012. doi: 10.1007/s10551-011-0948-0. [Cited on page 5.]

Y. Liu and V. Shankar. The dynamic impact of product-harm crises on brand preference and advertising effectiveness: An empirical analysis of the automobile industry. Management Science, 61(10):2514–2535, 2015. [Cited on page 1.]

P. Lloyd-Smith and H. An. Are corporate social responsibility and advertising complements or substitutes in producing firm reputation? Applied Economics, 51(21):2275–2288, 2019. [Cited on page 5.]

O. Lobel. Exit, voice & innovation: How human capital policy impacts equality (& how inequality hurts growth). Hous. L. Rev., 57:781, 2019. [Cited on page 1.]

O. Lobel. The contract thicket: Addition and supra addition in private law theory. Working paper, 2020. [Cited on pages 4 and 6.]

N. Maestas, K. J. Mullen, D. Powell, T. von Wachter, and J. B. Wenger. The value of working conditions in the united states and implications for the structure of wages. Technical Report 25204, National Bureau of Economic Research, October 2018. [Cited on pages 5 and 42.]

D. Mayzlin, Y. Dover, and J. Chevalier. Promotional reviews: An empirical investigation of online review manipulation. American Economic Review, 104(8):2421–55, August 2014. [Cited on page 39.]

R. C. McDevitt. Names and reputations: An empirical analysis. American Economic Journal: Microeconomics, 3(3):193–209, 2011. [Cited on page 1.]

T. Melo and A. Garrido-Morgado. Corporate reputation: A combination of social responsibility and industry. Corporate Social Responsibility and Environmental Management, 19(1):11–31, 2012. [Cited on page 5.]

D. Mertineit. Rare first circuit decision invalidating nda and overturning misappropriation verdict serves as a cautionary tale. https://www.lexology.com/library/detail.aspx?g=5c898b30-20d3-4d77-b897-794c2df7a84d, 2020. [Cited on page 43.]

E. W. Morrison. Employee voice and silence. Annu. Rev. Organ. Psychol. Organ. Behav., 1(1): 173–197, 2014. [Cited on page 5.]

P. K. Mowbray, A. Wilkinson, and H. H. Tse. An integrative review of employee voice: Identifying a common conceptualization and research agenda. International Journal of Management Reviews, 17(3):382–400, 2015. [Cited on page 5.]

A. Note. Law and economics look at contracts against public policy, 119 harv. L. Rev, 1445:1460, 2006. [Cited on page 4.]

R. M. Philip. Silence at our expense: Balancing safety and secrecy in non-disclosure agreements. Seton Hall. L. Rev., 33:845, 2002. [Cited on page 8.]

J. Prescott and E. Starr. Subjective beliefs about contract enforceability. Working Paper, 2021. [Cited on pages 3 and 41.]

A. Rees-Jones and K. T. Rozema. Price isn't everything: Behavioral response around changes in sin taxes. Technical report, National Bureau of Economic Research, 2019. [Cited on pages 3, 28, and 29.]

S. Sanga. Choice of law: an empirical analysis. Journal of Empirical Legal Studies, 11(4):894–928, 2014. [Cited on page 63.]

C. Shapiro. Premiums for high quality products as returns to reputations. The Quarterly Journal of Economics, 98(4):659–679, 1983. [Cited on page 1.]

J. Silver-Greenberg and N. Kitroeff. How bloomberg buys the silence of unhappy employees. www.nytimes.com/2020/03/02/business/michael-bloomberg-nda.html, 2020. [Cited on page 7.]

J. Sockin. Show me the amenity: Are higher-paying firms better all around? Technical Report 3957002, Social Science Research Network, November 2021. [Cited on pages 5 and 42.]

J. Sockin and M. Sockin. A pay scale of their own: Gender differences in variable pay. (3512598), December 2019. [Cited on page 36.]

J. Sockin and A. Sojourner. What's the inside scoop? challenges in the supply and demand for information on employers. (3596666), September 2020. [Cited on pages 1, 3, 5, 10, 11, 24, 39, and 42.]

I. Sorkin. Ranking firms using revealed preference. The Quarterly Journal of Economics, 133(3): 1331–1393, 2018. [Cited on pages 5 and 42.]

E. Starr, J. Frake, and R. Agarwal. Mobility constraint externalities. Organization Science, 30(5): 961–980, 2019. [Cited on page 5.]

E. Starr, J. J. Prescott, and N. Bishara. Behavioral effects of (unenforceable) contracts. Journal of Law, Economics, and Organization, 2020. [Cited on page 41.]

E. Starr, J. J. Prescott, and N. Bishara. Noncompete agreements in the u.s. labor force. Journal of Law and Economics, 2021. [Cited on pages 4 and 6.]

P. Sullivan and T. To. Search and nonwage job characteristics. Journal of Human Resources, 49 (2):472–507, 2014. [Cited on pages 5 and 42.]

S. Tadelis. What's in a name? reputation as a tradeable asset. American Economic Review, 89(3): 548–563, 1999. [Cited on page 5.]

N. Tiku, R. Albergotti, and G. De Vynck. California judge rules google?s confidentiality agreements break labor law. https://www.washingtonpost.com/technology/2022/01/14/google-nda-illegal-california/, 2021. [Cited on page 43.]

D. B. Turban and D. M. Cable. Firm reputation and applicant pool characteristics. Journal of Organizational Behavior, 24(6):733–751, 2003. [Cited on page 1.]

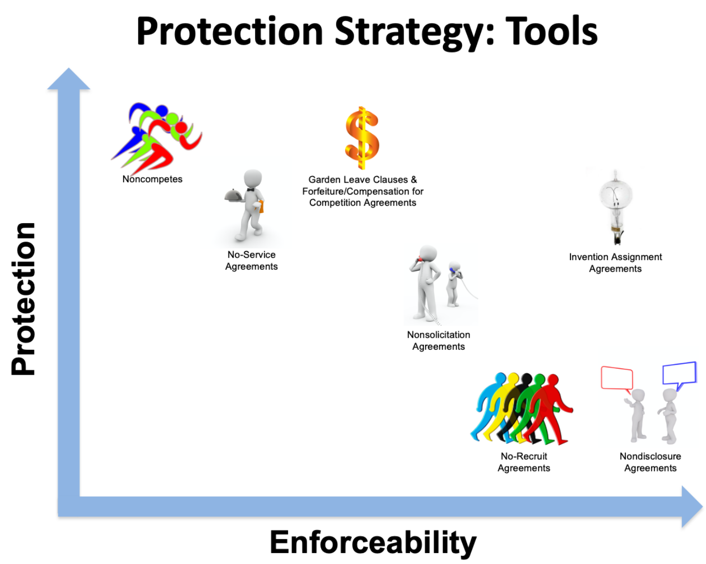
G. Usenheimer, N. Merola, and A. Dana. #metoo inspires legislative changes across the united states. https://www.lexology.com/library/detail.aspx?g=4470934c-7a2f-4ae9-86af-e670cd28799d, 2019. [Cited on page 9.]

T. Wilkinson-Ryan. The perverse consequences of disclosing standard terms. Cornell L. Rev., 103: 117, 2017. [Cited on page 41.]

J. Yang and J. Liu. Strengthening accountability for discrimination. Technical report, Economic Policy Institute, Washington, D.C., Jan. 2021. URL https://files.epi.org/pdf/218473.pdf. [Cited on page 4.]

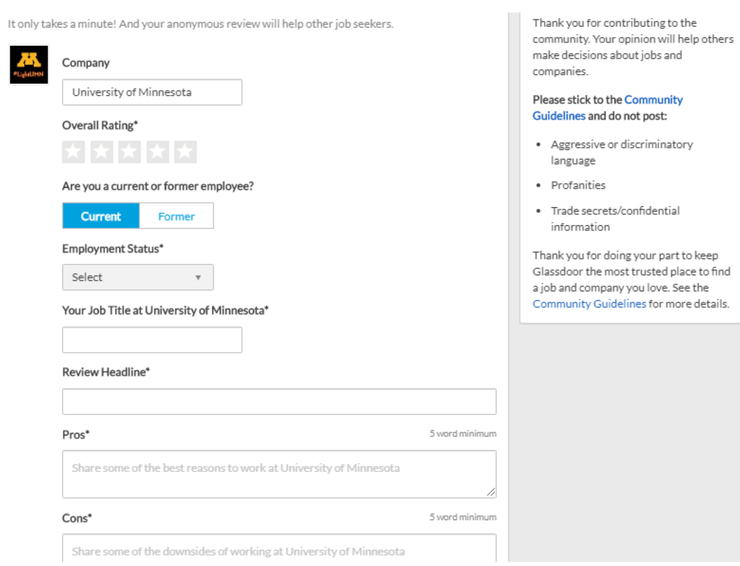
A Appendix: Additional Figures

Figure A1: Enforceability of NDAs and Related Restrictions



Notes: Figure available in Beck (2020).

Figure A2: Example of Blank Employer Review Form


It only takes a minute! And your anonymous review will help other job seekers.

Company

University of Minnesota

Overall Rating*

Are you a current or former employee?

Current Former

Employment Status*

Select

Your Job Title at University of Minnesota*

Review Headline*

Pros* 5 word minimum

Share some of the best reasons to work at University of Minnesota

Cons* 5 word minimum

Share some of the downsides of working at University of Minnesota

Thank you for contributing to the community. Your opinion will help others make decisions about jobs and companies.

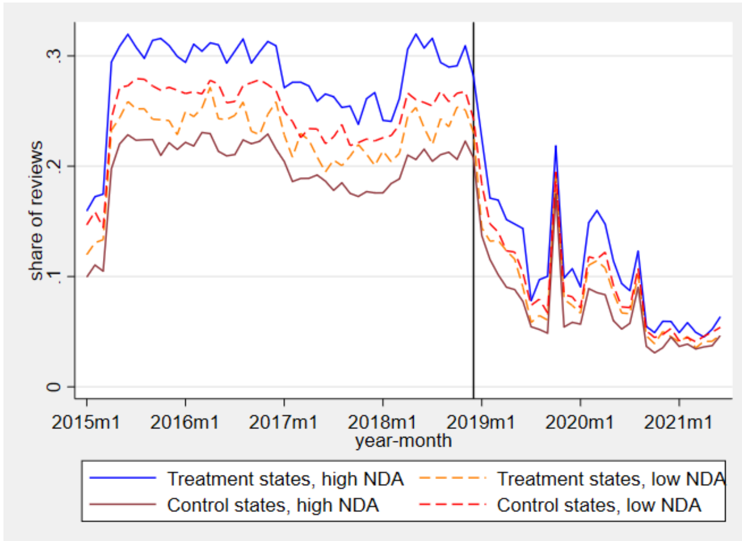
Please stick to the Community Guidelines and do not post:

- Aggressive or discriminatory language
- Profanities
- Trade secrets/confidential information

Thank you for doing your part to keep Glassdoor the most trusted place to find a job and company you love. See the Community Guidelines for more details.

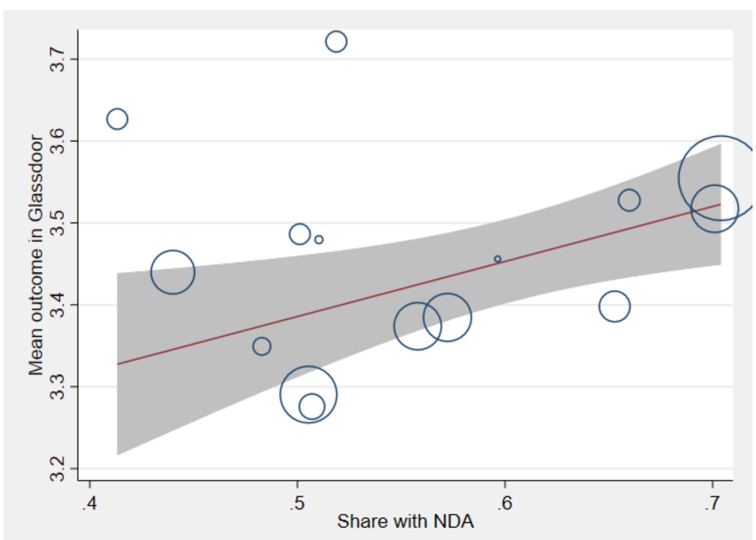
Notes: This figure is a screenshot of submitting an employer review for the University of Minnesota.

Figure A3: Structural Shift in Rates of Job Title Concealment Among 1–3 Stars Reviews



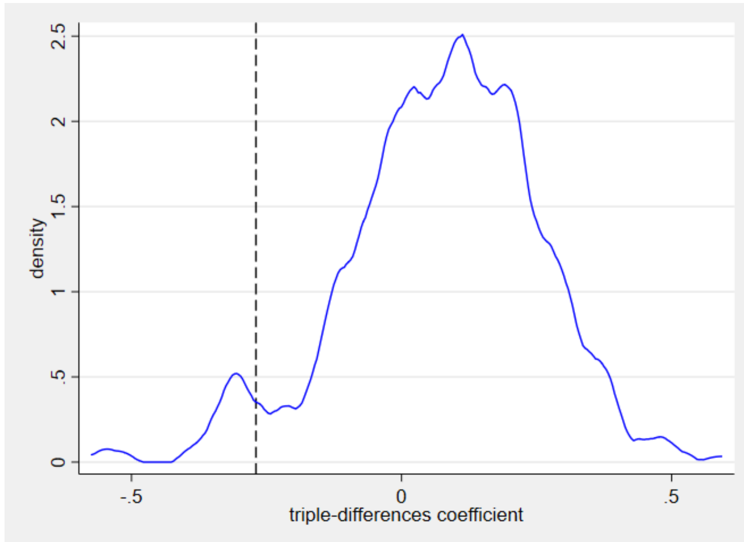
Notes: This figure plots the average rate of job title concealment among new reviews submitted each calendar month over the full sample period. Sample of reviews is partitioned into four subsets: treatment states (California, Illinois, and New Jersey) and control states (all others), as well as high NDA industries (above-average NDA intensity) and low NDA industries (below-average NDA intensity). Vertical line indicates the end of 2018, when there is a structural break in rates of job title concealment, likely associated with structural changes to the review submission platform. This structural shift occurs simultaneously with the enactment of California's legislation and a few months prior to the enactment of New Jersey's legislation. However, the change affected both treatment and control states and industries similarly.

Figure A4: NDA Incidence and Average Employer Rating by Industry



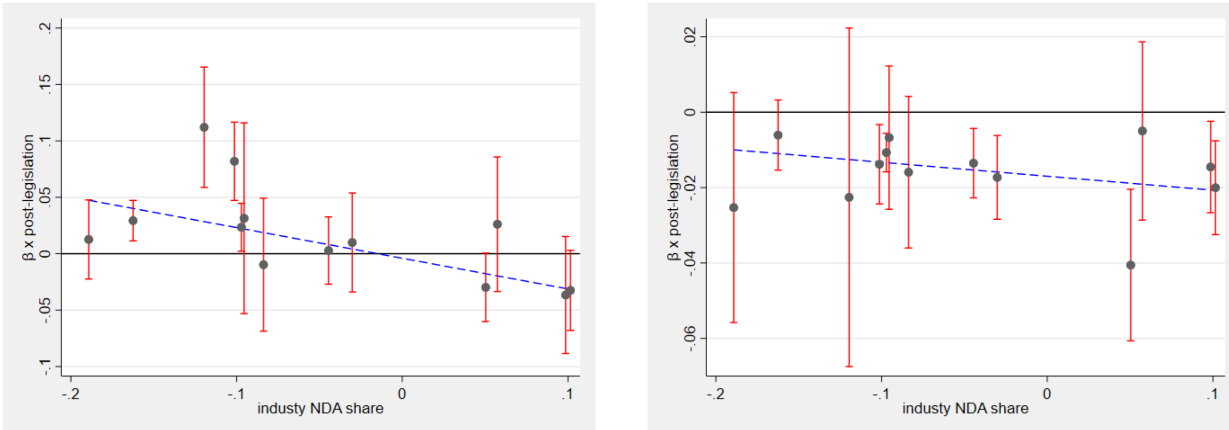
Notes: The figure illustrates the correlation between NDA incidence and average employee job satisfaction. Each dot represents an unique industry. Observations are weighted by the number of Glassdoor reviews within each industry. Industry-specific NDA incidence obtained from data through Payscale.com.

Figure A5: Distribution of Triple-Differences Estimates Under Randomized Inference



Notes: The figure illustrates the density of triple-differences coefficients when states are randomly assigned between treatment and control groups. There are three possible treatment periods: January 2019 (corresponding to California), March 2019 (corresponding to New Jersey), and January 2020 (corresponding to Illinois). We draw randomly from a uniform distribution to assign one of the fifty states or the District of Columbia to January 2019, a second of the fifty-one to March 2019, and a third to January 2020. The remaining forty-eight are assigned to the control group. We re-draw if the same state is assigned to two treatments. We then record the estimate from estimating a triple-differences specification. We repeat this procedure 500 times and plot the distribution. The dashed line indicates our main triple-differences estimate under the true treatment and control assignment.

Figure A6: Industry-Specific Difference-in-Differences Estimates for Overall Ratings

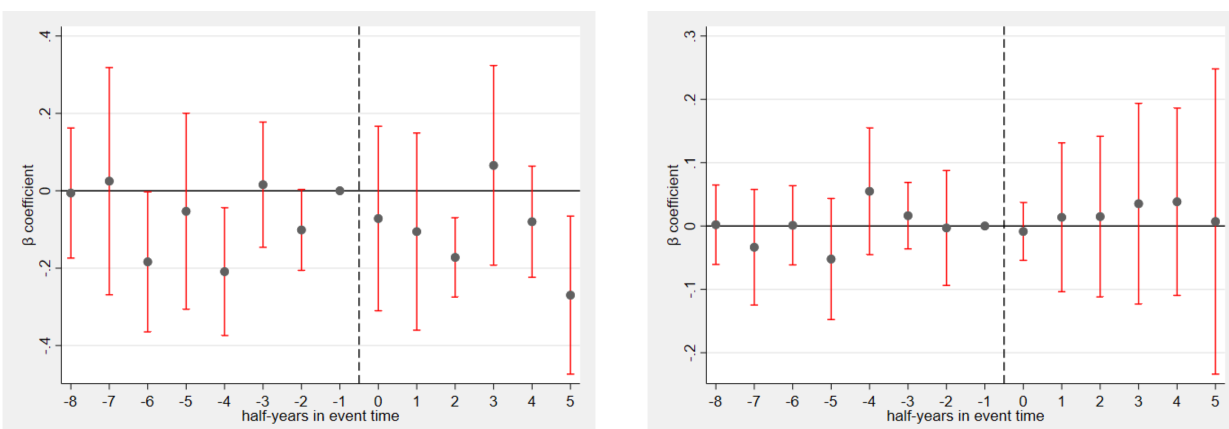


(a) Overall rating

(b) 1(Conceals job title | 1–3 stars)

Notes: These figures display difference-in-differences estimates from comparing CA-IL-NJ to all other states before and after NDAs are narrowed, separately for each industry. Regressions include firm-state and year-month fixed effects. Vertical red bars indicate a 95% confidence interval around each point estimate. Standard errors are clustered by state. Dashed blue line reflects linear lines of best fit through the point estimates, with industries weighted by their respective sample sizes. From left-to-right, the industries are: Construction, Accommodation and Food Services, Other Services, Arts and Entertainment, Retail Trade, Transportation and Warehousing, Real Estate, Health Care and Social Assistance, Manufacturing, Information, Utilities, Finance and Insurance, and Professional, Scientific and Technical Services. Excluded are two industries—Agriculture and Mining—for which standard errors are particularly large due to thin samples.

Figure A7: Narrowing NDAs and Labor Market Outcomes in the CPS, Dynamic Responses



(a) Log wage

(b) 1(same employer)

Notes: The figures present dynamic estimates for the how log wages (panel a) and the probability that workers switch employers (panel b) in the Current Population Survey evolve after NDAs are narrowed. The sample period is 2015–2021 and point estimates are relative to the calendar half-year before the legislation goes into effect. Samples are restricted to currently employed workers. Regressions are weighted using *earnwt* and *wtfinl* for panels (a) and (b), respectively. Standard errors are two-way clustered by industry and state. Red vertical bars indicate 95% confidence intervals around each point estimate.

B Appendix: Additional Tables

Table B1: Summary Statistics for Dependent Variables

Dependent variable	Reviews (millions)	Mean	Median	Standard deviation	5th percentile	95th percentile
Overall rating	3.89	3.49	4.00	1.41	1.00	5.00
Career opportunities	3.45	3.28	3.00	1.46	1.00	5.00
Compensation and benefits	3.45	3.36	4.00	1.36	1.00	5.00
Culture and values	3.43	3.44	4.00	1.52	1.00	5.00
Senior leadership	3.41	3.15	3.00	1.55	1.00	5.00
Work-life balance	3.45	3.38	4.00	1.44	1.00	5.00
Conceal job title | 1–3 stars	1.72	0.16	0.00	0.37	0.00	1.00
Would refer a friend to firm	3.31	0.63	1.00	0.48	0.00	1.00
Positive business outlook	3.18	0.54	1.00	0.50	0.00	1.00
Approve of CEO	2.66	0.52	1.00	0.50	0.00	1.00
Mentions harassment in review	3.89	0.02	0.00	0.14	0.00	0.00
Offers management advice	3.89	0.56	1.00	0.50	0.00	1.00
Log length of review text	3.89	5.28	5.15	0.87	4.14	6.88
Log length of pros section	3.89	4.42	4.23	0.87	3.30	6.09
Log length of cons section	3.89	4.54	4.36	1.00	3.26	6.44
Pros share of review text	3.89	0.48	0.48	0.20	0.13	0.82

Notes: Table provides summary statistics (mean, median, standard deviation, fifth percentile, and ninety-fifth percentile) for each of the review-level dependent variables.

Table B2: Narrowing NDAs and Within-Firm Dispersion in Ratings

	CA-IL-NJ		All states	
	Standard deviation	Interquartile range	Standard deviation	Interquartile range
Narrowed NDAs x NDA intensity	0.158**	0.228*	0.207***	0.311***
	(0.059)	(0.120)	(0.044)	(0.073)
Dependent variable mean	1.16	1.70	1.16	1.70
Mean reviews per firm-state-half	8.92	8.92	7.98	7.98
Observations	54831	54831	250790	250790
Adjusted R^2	0.09	0.09	0.09	0.08
Industry-Year-Half FE			✓	✓
State-Year-Half FE	✓	✓	✓	✓
Firm-State FE	✓	✓	✓	✓

Notes: The table implements triple-difference models for estimating the causal effect narrowing NDAs has on dispersion in overall ratings within firms. Each observation is the standard deviation or interquartile range of overall ratings for each firm-state-half year. Regressions are clustered by industry in the former two columns and two-way clustered by industry and state in the latter two. Samples are restricted to firm-states that average at least two ratings per calendar half year. Significance levels: * 10%, ** 5%, *** 1%.

Table B3: Narrowing NDAs and Alternative Outcomes for Employee Sentiment

	Star ratings					Indicators		
	Culture and values	Senior mgmt.	Career opp.	Comp. and benefits	Work-life balance	Would refer a friend to firm	Positive business outlook	Approve of CEO
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Narrowed NDAs x NDA intensity	-0.201***	-0.292***	-0.201**	-0.228***	-0.183***	-0.099***	-0.047*	-0.045**
	(0.058)	(0.070)	(0.076)	(0.064)	(0.034)	(0.013)	(0.024)	(0.015)
Dependent variable mean	3.434	3.136	3.284	3.364	3.370	0.628	0.534	0.520
Observations	3221143	3195791	3234875	3232153	3236662	3100240	2975550	2520492
Adjusted R^2	0.16	0.16	0.15	0.19	0.15	0.11	0.13	0.15
Industry-Year-Month FE	✓	✓	✓	✓	✓	✓	✓	✓
State-Year-Month FE	✓	✓	✓	✓	✓	✓	✓	✓
Firm-State FE	✓	✓	✓	✓	✓	✓	✓	✓

Notes: The table conveys how other outcomes related to newly submitted employer reviews changed, according to the triple-differences specification, following the narrowing of NDAs. The dependent variable in each regression is listed as the header of each column. Regressions include firm-state, industry-year-month, and state-year-month fixed effects. Standard errors are two-way clustered by industry and state. Significance levels: * 10%, ** 5%, *** 1%.

Table B4: Narrowing NDAs and Arrival of New Reviews

	Full US Within High NDA Ind.		Within CA-IL-NJ High vs. Low NDA Ind.		Triple Difference	
	(1)	(2)	(3)	(4)	(5)	(6)
CA-IL-NJ	2007*					
	(1053)					
Narrowed NDAs	1065	823*			610	
	(704)	(428)			(348)	
Narrowed NDAs x NDA intensity			4279	4374	2711	2714
			(4137)	(4324)	(1808)	(2127)
Dependent variable mean	646	646	1465	1465	392	392
Observations	2894	2894	585	585	9919	9919
Adjusted R^2	0.26	0.35	0.58	0.59	0.87	0.91
Half-Year FE	✓	✓	✓			
Industry FE	✓		✓	✓		
State FE		✓	✓	✓		
Industry-Half-Year FE					✓	✓
State-Industry FE					✓	✓
State-Half-Year FE				✓		✓

Notes: The dependent variable is the number of reviews submitted within an industry–state–half year. Regressions in first two columns are clustered by state, next two columns clustered by industry, and the final two columns clustered by industry and state. Significance levels: * 10%, ** 5%, *** 1%.

Table B5: Narrowing NDAs and Filings with Government Agencies

	EEOC			OSHA
	Total	Females	Males	
Narrowed NDAs	0.229*	0.151	0.290***	
	(0.135)	(0.113)	(0.096)	
Narrowed NDAs x NDA intensity				0.592
				(0.405)
Dependent variable mean	200.33	143.87	34.15	16.70
Observations	306	300	306	3992
Psuedo-R^2	0.95	0.94	0.82	0.89
Year FE	✓	✓	✓	
State FE	✓	✓	✓	
State-Industry FE				✓
Industry-Year FE				✓
State-Year FE				✓

Notes: Notes: The dependent variables are the number of EEOC charges filed alleging sexual harassment and the number of OSHA complaint-driven inspections. The EEOC data are at the state-year level, while the OSHA data are at the state-industry-year level. In both cases, fully saturated poisson fixed effects models are estimated. The sample period is 2015–2020 for EEOC and 2015–2021 for OSHA. Standard errors are two-way clustered by industry and state. Significance levels: * 10%, ** 5%, *** 1%.

Table B6: Narrowing NDAs and Dispersion in Firm Ratings using Interquartile Range

	Full US Within High NDA Ind.			Within CA-IL-NJ High vs. Low NDA Ind.		Triple Difference		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
CA-IL-NJ	-0.096***							
	(0.035)							
Narrowed NDAs	0.182***	0.193***	0.094***			0.084***		
	(0.045)	(0.050)	(0.024)			(0.031)		
Narrowed NDAs x NDA intensity				0.896*	0.849	0.954***	0.920***	0.469**
				(0.462)	(0.504)	(0.331)	(0.347)	(0.189)
Dependent variable mean	2.129	2.129	2.191	2.077	2.080	2.042	2.040	2.078
Observations	1534	1534	2652	468	585	4849	4823	9675
Adjusted R^2	0.18	0.26	0.25	0.19	0.20	0.47	0.47	0.53
Half-Year FE	✓	✓	✓	✓	✓			
State FE		✓	✓	✓	✓			
Industry-Half-Year FE						✓	✓	✓
State-Industry FE						✓	✓	✓
State-Half-Year FE							✓	✓
State-Industry 50+ firms on average	✓	✓		✓		✓	✓	
Weighted by average firm count			✓		✓			✓

Notes: The dependent variable in each regression is the interquartile range of the (mean) firm rating among firms within an industry–state–half year. Regressions in first three columns are clustered by state, the next two by industry, and the final three by industry and state. Significance levels: * 10%, ** 5%, *** 1%.

Table B7: Stacked Designs, Weighting, and Alternative Measures of NDA Intensity

Dependent variable	Industry stacked	Industry equally weighted	Occupation	Industry x occupation	Worker-level survey of employers	Firm-level survey
Overall rating	-0.274***	-0.384**	-0.183*	-0.153**	-0.053***	-0.163
	(0.080)	(0.168)	(0.098)	(0.062)	(0.014)	(0.114)
N	9913504	3654296	2577156	2001473	2067353	149994
Adjusted R^2	0.23	0.23	0.16	0.16	0.13	0.12
1(Conceals job title | 1–3 stars)	-0.058***	-0.082**	–	–	-0.003	0.021
	(0.015)	(0.030)	–	–	(0.005)	(0.015)
N	4395474	1569510	–	–	919860	69556
Adjusted R^2	0.23	0.24	–	–	0.10	0.12

Notes: The table conveys the triple-difference estimates when alternative regression specifications or alternative measures of NDA intensity are used in lieu of the benchmark model with industry-level NDA intensity. The dependent variable in each regression is employee overall star rating. The "Industry stacked" model replicates the control sample thrice–for each treatment state–incorporating firm-state, state-year-month, industry-year-month, and treatment-state sample fixed effects. Standard errors are two-way clustered by industry and state. The "Industry equally weighted" model weights each review by $1/N_{\iota(k)st}$ such that each industry-state-year receives equal weight, where the sample is restricted to industry–state pairings that receive on average at least 100 reviews annually. Standard errors are two-way clustered by industry and state. The "industry x occupation" model use industry–occupation level NDA intensity from Payscale data in lieu of industry level incorporating firm-state, occupation-state, occupation-industry, occupation-year-month, industry-year-month, and state-year-month fixed effects and restricting the sample to industry–occupation pairs with at least twenty observations in the Payscale survey. Standard errors are two-way clustered by industry-occupation and state. The "occupation" model use occupation level NDA intensity from Payscale data in lieu of industry level incorporating firm-state, occupation-state, occupation-year-month, and state-year-month fixed effects and restricting the sample to occupations with at least twenty observations in the Payscale survey. Standard errors are two-way clustered by occupation and state. The "Employer specific worker survey" uses data from the individual-level survey underlying the industry-level estimates (further described in Balasubramanian et al. (2021)) in which individuals were asked if they had an NDA. We use this worker-level, employer-specific NDA measure in lieu of industry level by first matching observations in the Payscale survey to firms and calculating the share of respondents per firm with an NDA, restricting the sample to firms with at least one observation in the Payscale survey and incorporating firm-state, state-year-month, and industry-year-month fixed effects. Standard errors are two-way clustered by firm and state. The "Employer specific firm survey" uses data from a separate firm-level survey deployed by Payscale (also described in Balasubramanian et al. (2021)) in which firms are asked whether all, some, or none of the firm's employees are covered by NDAs. With this measure we create an indicator that is active for the first two and not active for the latter—in lieu of industry level NDA intensity and incorporating firm-state, state-year-month, and industry-year-month fixed effects. Standard errors are two-way clustered by firm and state. Significance levels: * 10%, ** 5%, *** 1%.

Table B8: Alternative Choice of Control States

Dependent variable	All states	Neighbor states	High coverage states	Weaker legislation states
Overall rating	-0.270***	-0.264**	-0.235**	-0.189*
	(0.079)	(0.102)	(0.082)	(0.093)
Observations	3654296	1632233	2521647	1490471
Adjusted R^2	0.15	0.16	0.15	0.16
1(Conceals job title | 1–3 stars)	-0.060***	-0.056*	-0.039*	-0.027
	(0.015)	(0.031)	(0.019)	(0.028)
Observations	1569510	698243	1079657	627150
Adjusted R^2	0.13	0.13	0.13	0.14
Standard deviation firm ratings (weighted)	0.061**	0.084**	0.091***	0.104***
	(0.031)	(0.033)	(0.031)	(0.035)
Observations	9401	2667	2519	1880
Adjusted R^2	0.62	0.72	0.81	0.75
IQR firm ratings (weighted)	0.469**	0.389*	0.585**	0.614***
	(0.189)	(0.213)	(0.225)	(0.221)
Observations	9675	2700	2533	1914
Adjusted R^2	0.53	0.58	0.63	0.59

Notes: The table illustrates the robustness of our main triple-differences estimates to the choice of control sample. The "All states" sample includes the other forty-seven states and the District of Columbia. The "Neighbor states" reflects the eleven states (Arizona, Delaware, Indiana, Iowa, Kentucky, Missouri, Nevada, New York, Oregon, Pennsylvania, and Wisconsin) that share a contiguous border with California, Illinois, or New Jersey. The "High coverage states" refer to the ten states (other than California, Illinois and New Jersey) that represent at least 2.5% of the review sample (Florida, Georgia, Massachusetts, New York, North Carolina, Ohio, Pennsylvania, Texas, Virginia, and Washington). And "Weaker legislation states" refers to the seven states that implemented more narrowly-focused legislation regarding the use of non-disclosures around the same time (Maryland, New York, Oregon, Tennessee, Virginia, Vermont, and Washington). These laws were weaker than those studied here because they typically covered only NDAs in the context of sexual harassment. See Johnson et al. (2019) for details. Standard errors are two-way clustered by industry and state. Significance levels: * 10%, ** 5%, *** 1%.

Table B9: Alternative Choice of Standard Error Clustering

	Clustering of standard errors				
Dependent variable	Industry and state	State	Industry cross state	Firm and state	Wild cluster bootstrap
Overall rating	-0.270***	-0.270***	-0.270***	-0.270***	-0.273*
	(0.079)	(0.086)	(0.047)	(0.083)	(0.107)
1(Conceals job title | 1–3 stars)	-0.060***	-0.060***	-0.060***	-0.060***	-0.058*
	(0.015)	(0.019)	(0.016)	(0.016)	(0.019)
Standard deviation firm ratings (weighted)	–	0.061***	0.061***	–	0.061
		(0.017)	(0.031)		(0.035)
IQR firm ratings (weighted)	0.469***	0.469***	0.469**	–	0.469**
	(0.147)	(0.133)	(0.189)		(0.214)

Notes: The table re-estimates each of the main difference-in-differences specifications (overall rating of reviews, job title concealment among the most negative reviews, and the dispersion in firm ratings) under different clustering methods for the standard errors. The baseline clustering method is two-way clustering by industry and state for overall rating and job title concealment, and, due to convergence issues, industry cross state for the standard deviation of firm ratings. For ease of implementing wild cluster bootstrapping due to large sample sizes for overall rating and job title concealment, we relax the time-related fixed effects to industry-year and state-year. For wild cluster bootstrapping, we conduct 500 replications. Significance levels: * 10%, ** 5%, *** 1%.

Table B10: Effect on Average Rating Accounting for Differences in Wage Growth

	Level of aggregation for mean earnings				
	Industry x State	Occupation x State	Industry x Occupation x State	Firm x State	Firm x Occupation x State
Narrowed NDAs x NDA intensity	-0.301***	-0.312***	-0.328***	-0.283***	-0.256***
	(0.073)	(0.066)	(0.072)	(0.069)	(0.052)
Mean log earnings	-0.033	0.174***	0.156***	0.169***	0.175***
	(0.043)	(0.031)	(0.029)	(0.013)	(0.027)
Observations	2958774	1840988	1790316	2431784	980746

Notes: The table conveys the triple-difference estimates when average pay within a labor market is incorporated into the model. Mean log earnings is calculated using Glassdoor pay data by calendar half-year at the level of aggregation detailed in the header of each column. Sample excludes reviews from the first half of 2021 because our Glassdoor pay dataset extends only as far as 2020. The dependent variable in each regression is employee overall star rating. NDA intensity reflects the benchmark industry level. Samples for Columns 2, 3, and 5 are necessarily restricted to reviews for which job title is available. Standard errors are two-way clustered by industry and state. Significance levels: * 10%, ** 5%, *** 1%.

Table B11: Imputation of Missing Location

	Overall rating				
	(1)	(2)	(3)	(4)	(5)
Narrowed NDAs x NDA intensity	-0.270***	-0.236***	-0.205***	-0.206***	-0.213***
	(0.079)	(0.074)	(0.069)	(0.069)	(0.071)
Threshold for including missing location	none	100%	75%	50%	25%
Dependent variable mean	3.478	3.470	3.468	3.467	3.461
Observations	3654296	3976269	4482519	4926563	5511901
Adjusted R^2	0.15	0.15	0.15	0.15	0.15
Industry-Year-Month FE	✓	✓	✓	✓	✓
State-Year-Month FE	✓	✓	✓	✓	✓
Firm-State FE	✓	✓	✓	✓	✓

Notes: The table explores whether the exclusion of reviews without the location left blank is driving the reduction in average ratings following the passage of these laws. The dependent variable in each regression is employee overall star rating. Regressions include firm-state, industry-year-month, state-year-month, a location left blank dummy–year–month fixed effects. Standard errors are two-way clustered by industry and state. Significance levels: * 10%, ** 5%, *** 1%.

Table B12: Isolating the Treatment Effect of Narrowing NDAs by State

	Lone Treatment State		
	California	Illinois	New Jersey
	(1)	(2)	(3)
Narrowed NDAs x NDA intensity	-0.369***	-0.136***	0.043
	(0.032)	(0.033)	(0.038)
Observations	3379919	3018267	2927558
Adjusted R^2	0.15	0.15	0.15
Industry-Year-Month FE	✓	✓	✓
State-Year-Month FE	✓	✓	✓
Firm-State FE	✓	✓	✓

Notes: The table illustrates the heterogeneity by treatment state underlying our main triple-differences estimate. For each column, we exclude two treatment states entirely and estimate our main triple-difference with the lone treatment states listed in each column sub-header. For column (1), Illinois and New Jersey are dropped; for column (2), California and New Jersey; for column (3), California and Illinois. Standard errors are two-way clustered by industry and state. Significance levels: * 10%, ** 5%, *** 1%.

Table B13: Effect on Overall Rating by NDA Intensity Following Weinstein Scandal

	CA only		CA, IL, NJ		Triple
	3 mths	12 mths	3 mths	12 mths	difference
After Weinstein Scandal x NDA intensity	-0.261**	-0.006	-0.279*	-0.077	-0.040
	(0.095)	(0.049)	(0.066)	(0.055)	(0.087)
Narrowed NDAs x NDA intensity					-0.270***
					(0.079)
Observations	86189	133928	132314	203849	3654296
Adjusted R^2	0.20	0.19	0.19	0.18	0.15
Year-Month FE	✓	✓	✓	✓	
Firm FE	✓	✓	✓	✓	
Industry-Year-Month FE					✓
State-Year-Month FE					✓
Firm-State FE					✓

Notes: The table reflects how overall ratings evolve in California, Illinois, and New Jersey following the public revelation of the Harvey Weinstein scandal on October 5th, 2017. The pre-period for each specification is the twelve calendar months preceding this date. Short-term and long-term effects following this event are estimated using post-periods of three and twelve calendar months, respectively. Standard errors for the former two columns are clustered by industry, the latter two columns by industry and state. Significance levels: * 10%, ** 5%, *** 1%.

Table B14: Narrowing NDAs and the extent of Review Planting

	Growth threshold for identifying sock puppetry		
	100%	50%	25%
Narrowed NDAs x NDA intensity	0.000	-0.004	-0.044**
	(0.015)	(0.011)	(0.017)
Dependent variable mean	0.073	0.210	0.355
Observations	3654296	3654296	3654296
Adjusted R^2	0.43	0.39	0.37
Industry-Year-Month FE	✓	✓	✓
State-Year-Month FE	✓	✓	✓
Firm-State FE	✓	✓	✓

Notes: The table explores whether the presence of reviews identified as possible sock puppetry changes following the passage of these laws. For each firm in each year-month, we calculate the log change in reviews relative to the three months prior (g_{kt}^B) and the three months after (g_{kt}^A). The x% percent cutoff refers to firm-year-months in which $g_{kt}^B \geq x\%$ and $g_{kt}^A \geq x\%$. The dependent variable for each specification reflects an indicator variable for satisfying this criteria. Regressions include firm-state, industry-year-month, and state-year-month fixed effects. Standard errors are two-way clustered by state and industry. Significance levels: * 10%, ** 5%, *** 1%.

Table B15: Accounting for Possible Review Farming in Submission of Reviews

	Overall rating			
	(1)	(2)	(3)	(4)
Narrowed NDAs x NDA intensity	-0.270***	-0.278***	-0.259***	-0.249***
	(0.079)	(0.075)	(0.071)	(0.060)
Growth threshold for excluding reviews	none	100%	50%	25%
Dependent variable mean	3.478	3.431	3.418	3.419
Observations	3654296	3374684	2850402	2316669
Adjusted R^2	0.15	0.14	0.13	0.13
Industry-Year-Month FE	✓	✓	✓	✓
State-Year-Month FE	✓	✓	✓	✓
Firm-State FE	✓	✓	✓	✓

Notes: The table explores whether the presence of reviews identified as possible sock puppetry are driving the reduction in average ratings following the passage of these laws. For each firm in each year-month, we calculate the log change in reviews relative to the three months prior (g_{kt}^{B}) and the three months after (g_{kt}^{A}). The x% percent cutoff refers to firm-year-months in which $g_{kt}^{B} \geq x\%$ and $g_{kt}^{A} \geq x\%$. The sample for each specification *excludes* reviews from all firm-year-months satisfying this criteria. The dependent variable in each regression is employee overall star rating. Regressions include firm-state, industry-year-month, and state-year-month fixed effects. Standard errors are two-way clustered by state and industry. Significance levels: * 10%, ** 5%, *** 1%.

Table B16: Narrowing NDAs and Outcomes in the Current Population Survey

	Log wage	1(Same employer)
Narrowed NDAs x NDA intensity	-0.019	0.018
	(0.040)	(0.024)
Dependent variable mean	2.757	0.976
Observations	489762	416620
Adjusted R^2	0.28	0.01
Industry-Year-Month FE	✓	✓
State-Industry FE	✓	✓
State-Year-Month FE	✓	✓

Notes: The table implements triple-differences models estimating the causal effect narrowing NDAs had on the average wage and the probability workers switch employers using monthly data spanning 2015–2021 from the Current Population Survey. Regressions are weighted using representative weights available in the CPS. Samples are restricted to currently employed workers. Standard errors are two-way clustered by industry and state. Standard errors are two-way clustered by state and industry. Significance levels: * 10%, ** 5%, *** 1%.

C Heterogeneity in Ratings

Our heterogeneity analyses are driven by both theoretical and practical concerns. From a theoretical perspective, NDAs threaten legal and financial costs on workers if they speak out and share negative information. Accordingly, narrowing NDAs is likely to have the strongest effects on ratings from workers who have more negative information to share and who experience larger reductions in the probability of negative consequences of violating an NDA. Two hypotheses follow. First, while both current and former employees would face legal risk from violating an NDA, current employees face more substantial potential retaliation by the firm, given that a current worker is still employed there. Therefore, the laws' new protections against firm retaliation for violating an NDA may have greater bite for current employees. Second, given that 84.1% of the sexual harassment charges filed with the EEOC in 2018 are from women, they likely have more negative information to share.

Table C1 shows the results of our main triple-differences specification, but allowing for heterogeneous estimates within various sample partitions. Columns (1) and (2) split the sample by whether the individual is a current or former employee and finds that while both provide more negative reviews after NDAs are narrowed, current employees increase provision of negative information more than former employees, though the difference is not precise. Columns (3) and (4) split the sample into workers who had or currently have a short tenure (at most two years) with the firm or a long tenure (more than two years). The negative effects we observed are driven more by shorter-tenure employees. Column (5) and (6) split the sample by gender. Contrary to our expectation, men and women similarly provide more-negative reviews. Many respondents do not report their gender, perhaps to purposefully obscure their identity, which adds noise to these cuts of the data. Nevertheless, the results do not suggest that narrowing NDAs causes women to share more negative information than men.

We also examine three dimensions of firm heterogeneity. First, one way that firms can potentially avoid the policies that weakened NDAs is by using choice of law and forum provisions that stipulate that in the event of contract breach a different state's law be applied (Sanga, 2014; Coyle, 2020). While we do not know what choice of law/forum provision each firm has in its employment contract, we design our empirical specification based on the assumption that single-state employers have chosen the laws of the state in which they operate, while multi-state employes could choose myriad state laws. We divide firms based on whether their employer reviews on Glassdoor stem from a single state, or come from multiple states. Columns (7) and (8) reports the results, splitting the sample by whether the firm is a single-state or multi-state employer. Consistent with expectation, we find that the negative reviews stemming from narrowing NDAs are driven more by firms operating in a single state, though the effects are not statistically distinguishable from that observed among multiple-state firms.

Second, we allow for different effects for small and large firms, as workers employed by smaller firms likely face greater risk of retaliation from being less able to blend in among a pool of coworkers. We define a small (large) firm as an employer whose size is below (above) the sample median. Columns (9) and (10) reveal that while the effect is negative for both, it is larger among small firms (though not statistically distinguishable), consistent with workers at smaller firms feeling less burdened by retaliatory risk following the passage of these laws.

Third, an important question is whether the decline is driven by bad-reputation firms receiving worse reviews, or good-reputation firms whose reputations were inflated by the use of NDAs. It is possible that the returns to reputation are non-linear, such that firms with 'good' reputations have stronger incentives to keep negative information from coming out. We classify firm-state pairs by whether the average rating among reviews submitted in 2018—the year prior to enactment of either of the three laws—was above or below average, and re-estimate our triple-differences specification on each sub-sample. The former we label as "high rated" firms and the latter "low rated" ones. In Columns (11) and (12), we find that both low-rated and high-rated firms receive more-negative reviews after NDAs are narrowed, and the difference in the declines is not statistically significant.

Table C1: Heterogeneity in Narrowing NDAs and Employees' Overall Ratings of Firms

	Worker characteristic						Firm characteristic					
	Current	Former	Short tenure	Long tenure	Male	Female	Operates one state	Operates many states	Small	Large	Low rated	High rated
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Narrowed NDAs x NDA intensity	-0.255***	-0.194***	-0.273***	-0.199***	-0.286**	-0.231*	-0.409**	-0.253***	-0.279**	-0.205**	-0.239***	-0.299**
	(0.073)	(0.037)	(0.060)	(0.062)	(0.105)	(0.117)	(0.160)	(0.067)	(0.125)	(0.077)	(0.032)	(0.137)
Dependent variable mean	3.87	3.02	3.47	3.46	3.45	3.33	3.60	3.46	3.54	3.42	3.18	3.75
Observations	1923702	1621254	1709088	1088438	917549	775468	394951	3259222	1705491	1904800	1481629	1541990
Adjusted R^2	0.19	0.16	0.17	0.17	0.16	0.15	0.20	0.14	0.18	0.11	0.11	0.13
P-value for test of difference		0.504		0.042		0.468		0.171		0.645		0.624
Industry-Year-Month FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
State-Year-Month FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Firm-State FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Notes: The table investigates whether the effect on average overall ratings following the passage of these three laws under the triple-differences specification differs among various subsets of reviews, partitioned according to either worker or firm characteristics. Short (long) tenure employees refer to workers with at most (least) two years of work experience at the firm. The number of states the firm operates in is determined by calculating the number of unique states from which there is an employee review in the Glassdoor data. Volunteers are not asked to reveal their gender when submitting an employer review, but for a subset of respondents, gender is obtained through other aspects of the platform, such as a user profile. Low (high) rated firms reflect employers for which their average overall ratings in 2018 are above (below) average. Regressions include firm-state, industry-year-month, and state-year-month fixed effects. Standard errors are two-way clustered by industry and state. Significance levels: * 10%, ** 5%, *** 1%.