Accusing and Believing in Equilibrium: Evidence from #MeToo*

Anya Marchenko† Brown University

Bobak Pakzad-Hurson‡ Brown University

March 26, 2026

Abstract

Social movements that encourage reporting of wrongdoing may also induce unfounded accusations. We study how #MeToo—a social movement that encouraged reporting of sexual harassment—changed the strength of sexual harassment accusations and adjudicators' willingness to believe them. We estimate that #MeToo increased the probability of winning a sexual harassment complaint by 10.1 pp. This increase can reflect stronger complaints filed (selection) or more favorable adjudicators (direct treatment). We develop a theoretical framework to disentangle these effects, and use complaints filed before but resolved after #MeToo to estimate the direct treatment effect. We find #MeToo caused adjudicators to be more favorable to complainants and that complaints induced by #MeToo were stronger. However, there was significant gender heterogeneity: women induced to report brought stronger complaints, while men induced to report brought weaker complaints. In other words, #MeToo encouraged reporting of unfounded accusations only by men.

JEL Classification: J16, K42, J71.

*We thank Abi Adams, Laura Boudreau, Adamson Bryant, Sylvain Chassang, Olle Folke, Andy Foster, Ada Gonzalez-Torres, Joni Hersch, Peter Hull, Christine Jolls, Lorenzo Lagos, Martin Mattsson, Emily Nix, Matthew Pecenco, Johanna Rickne, Jonathan Roth, and Olga Stoddard for helpful comments. Abigail Flamer, Jacob Hirschhorn, and Maggie Jiang provided excellent research assistance. We are grateful for support from the Orlando Bravo Center for Economic Research. Marchenko acknowledges support from the National Science Foundation Graduate Research Fellowship and the Dr. Frances Price Harnish 1925 Fellowship.

†Marchenko: marchenko@brown.edu

‡Pakzad-Hurson: bph@brown.edu

1 Introduction

The goal of many social movements and campaigns is to increase the reporting of wrongdoing. For example, the "See Something, Say Something" campaign encourages the reporting of terrorism-related activity, Black Lives Matter the reporting of police misconduct, and Environmental, Social, and Governance (ESG) initiatives the disclosure of harmful corporate behavior. But there is debate about whether more reporting exposes legitimate wrongdoing or generates excessive accusations. We study this debate in the context of the #MeToo movement, one of the largest social movements of the twenty-first century, which marked a turning point in public attitudes towards sexual harassment. Starting in 2017, the movement spurred a flurry of sexual harassment and sexual assault accusations against high-profile individuals in media, politics, and academia. The goal was to decrease the costs of reporting sexual harassment (by decreasing stigma) and increase the benefits (by believing more complainants).

Public opinion diverged after #MeToo, with some believing it encouraged unfounded accusations. For example, the majority of people surveyed by NPR[1] thought #MeToo created accountability for perpetrators, but 40% thought it went too far:

> "respondents cited a rush to judgment, the prospect of unproven accusations ruining peoples' careers or reputations, and a bandwagon effect that may prompt some to claim sexual misconduct for behavior that doesn't quite rise to that level."

There are perhaps two chiefly relevant ways to measure the credibility of accusations made in the wake of the movement: First, are adjudicators more likely to believe them, all else equal? Second, are they "stronger" according to an objective metric? Following conventions in the justice system and prior literature (Lisak et al., 2010; Lonsway and Archambault, 2012), we proxy for the strength of an accusation by whether it is *founded* or *unfounded*. The FBI defines an accusation as being unfounded if it is either false (i.e., fabricated) or baseless (i.e., the incident is truthfully recounted but does not meet the legal threshold for sexual harassment). Therefore, our second question can be rephrased to ask whether accusations are more or less likely to be founded as a result of #MeToo.

We investigate these questions jointly, as they are potentially interdependent in equilibrium. If

[1] See https://www.npr.org/2018/10/31/662178315/on-metoo-americans-more-divided-by-party-than-gender

adjudicators are more likely to believe accusations, then more individuals may make accusations, and these induced accusations may or may not be more likely to be founded. On the other hand, if many more founded (unfounded) accusations are made, then adjudicators may be more (less) likely to believe them. To answer these questions, we first empirically estimate the impact of #MeToo on the probability that different populations of complainants win their cases. We then use a theoretical model to decompose these empirical estimates into two effects: a *direct treatment* effect ("Is the adjudicator more or less willing to believe a given complaint after #MeToo?") and a *selection into reporting* effect ("Are complaints induced by #MeToo more or less likely to be founded than prior complaints?").

Our data comprise sexual harassment complaints — such as inappropriate comments or requests for sexual favors — brought to state anti-discrimination agencies. These agencies serve as neutral third-party arbiters, offering a standardized process to resolve discrimination disputes. Following a complaint, the agency conducts an investigation into the merits of the claim, which may include visiting the employer site, conducting interviews, or reviewing personnel files. The agency then rules on whether discrimination occurred. These cases are civil as opposed to criminal, meaning that the burden of proof the complainant must reach in order to be "believed" is a preponderance of the evidence: a winning complaint is one where discrimination is assessed more likely to have occurred than not (i.e., more likely to be founded than unfounded).

Sexual harassment complaints are an ideal setting to study the joint impacts of the direct treatment and selection into reporting effects. First, because of the low evidentiary threshold, it is plausible for a fraction of accusations to be unfounded; with a higher burden of proof, all but the most substantiated accusations may not be made in equilibrium. Second, sexual harassment cases are frequently of a "he-said-she-said" nature and lack the forensic evidence available in criminal cases. Given the burden of proof, small changes in how adjudicators weigh such evidence can cause large changes in the overall win rate of complaints. Third, our setting lacks formal court procedures (e.g., rules on the consideration of character evidence) and therefore may give more leeway for adjudicators' beliefs to guide the final ruling.[2]

[2]Consider the case of Harvey Weinstein, a Hollywood producer who was accused of sexual assault—a criminal charge—by many women. In 2024, an appeals court overturned earlier convictions. The central disagreement was over the role of testimonial evidence in the original trial and whether it should count for his conviction. In a dissenting opinion (see https://www.nycourts.gov/ctapps/Decisions/2024/Apr24/24opn24-Decision.pdf), Judge Singas argued

We proceed in two parts. In the first part of the paper, we study the average impact of #MeToo on the outcomes of workplace sexual harassment complaints. To do so, we compile a new dataset of employment discrimination complaints by filing Freedom of Information Act requests with state anti-discrimination agencies. We receive a total of 192,000 usable complaints filed between 2010 and 2023.[3] We observe the basis of discrimination alleged, filing and resolution dates, the outcome, and in some states the complainant's name and gender.[4]

We use difference-in-differences to estimate how #MeToo changed the outcomes of sexual harassment complaints as compared to a control group of other discrimination complaints. Age, disability, nationality, race, religion, and retaliation discrimination complaints form our control group. Our main outcome is the probability that the complainant wins (i.e., there is probable cause to believe discrimination occurred). The win rate is a function of both the quality of evidence in the complaint as well as how favorable adjudicators are, conditional on a certain evidence level. An increase in the win rate can therefore reflect complaints filed with better evidence, more favorable adjudicators,[5] or both. Since this average impact of #MeToo combines the selection and direct treatment effects, we refer to it as the *combined effect*. We find that #MeToo had a combined effect of 10.1 percentage points, increasing the probability that sexual harassment complainants win by 66% over the control mean. We also use a triple difference specification to examine effects by gender and find the effect for women is positive and significant, but not statistically different from the effect for men. While this finding may be surprising given #MeToo mainly focused on female victims, it potentially masks substantial gender heterogeneity in both selection into reporting and direct treatment.

The second part of our paper develops a framework for comparing the credibility of complaints and adjudicator behavior before and after #MeToo, even when the evidence presented in a given

that the majority fails to give accurate weight to women's accusations, thereby advantaging Weinstein. She stated, "by whitewashing the facts to conform to a he-said/she-said narrative...this Court has continued a disturbing trend of overturning juries' guilty verdicts in cases involving sexual violence." The setting we study does not have such formal evidentiary rules, allowing for the possibility that testimonial evidence plays a strong role in determining final outcomes.

[3]The Equal Employment Opportunity Commission – the federal analogue of these state agencies – no longer provides information on outcomes of individual sexual harassment complaints via Freedom of Information Act requests, making the data in this paper the only detailed public record of both filings and outcomes of US workplace sexual harassment complaints post-#MeToo to our knowledge.

[4]The defendant is typically listed as an employer rather than an individual being accused, so we cannot consider heterogeneity by defendant gender.

[5]Other changes affecting the overall leniency of investigations, such as more favorable treatment of complaints by staffers or witnesses, are included in this "adjudicator" effect.

complaint is not directly observable in the data. We present a simple model in which each complaint is either founded or unfounded. The distribution of complaints can change over time; for example, if #MeToo deters sexual harassment, the share of founded complaints may change. Founded complaints are assumed to typically generate more evidence than unfounded ones, and adjudicators base their rulings on the amount of evidence present via a threshold rule: a complainant "wins" if and only if her complaint generates sufficient evidence. Therefore, founded cases win more often in expectation. This links win rates in the data and the underlying composition of founded versus unfounded complaints: first, holding fixed the share of founded allegations, a higher win rate after #MeToo indicates that adjudicators lower their evidentiary threshold and "believe" more allegations. Second, holding fixed adjudicators' evidentiary threshold, a sample with a higher win rate must have a higher share of founded complaints than a sample with a lower win rate.

We use our model findings to decompose the combined effect of #MeToo under plausible assumptions. The key intuition behind our decomposition is that the post-period includes two types of reporters: *always reporters* (a subset of reporters who are as likely to make a founded accusation as reporters pre-#MeToo) and *induced reporters* (the remaining reporters, whose complaints may be more or less likely to be founded). Importantly, because always reporters are neither positively nor negatively selected by definition, the impact of #MeToo on the win rate for always reporters reflects only the direct treatment effect on adjudicators. We prove that, under our assumptions, the combined effect can be written as a weighted average of the average treatment effects for these two groups, where the weights correspond to their shares in the post period.

To estimate the effect for always reporters, we exploit *overlap complaints* — cases filed before but resolved after #MeToo — which, by construction, could not have been filed in response to the movement. Since overlap complaints are mechanically not subject to selection, the treatment effect on these cases is the effect on always reporters (i.e., the direct treatment effect). Comparing this overlap effect to the combined effect reveals the direction of selection for induced reporters. Intuitively, if the overlap effect is smaller than the combined effect, induced reporters must be positively selected to make up the difference; if it is bigger, they must be negatively selected. The size of the selection effect, but not its sign, depends on the (latent) share of complainants who are always reporters. The bigger this share, the larger is the selection effect. We present a decomposition

which yields the relative size of the direct treatment versus selection effects as a function of the share of always reporters, and provide calibrations on this share from related studies.

We find that the effect on overlap complaints is 8.7 percentage points, smaller than the combined effect of 10.1 percentage points. This leads to two findings of economic interest. First, the effect on overlap complaints is positive. Through the lens of our model, this implies that adjudicators were more likely to believe accusations of sexual harassment post-#MeToo, holding the characteristics of the complaint fixed. Second, since the effect without selection is smaller than the combined effect, the effect on induced reporters must be positive and large to make up the difference. This implies that complaints induced to report by #MeToo are *positively* selected for foundedness.

However, these aggregate findings disguise interesting gender heterogeneity: the positive selection in new cases is driven by women. Our theoretical framework easily extends to within-gender comparisons (or other within-group comparisons), meaning that we can investigate the treatment and selection effects within gender analogously to our aggregate approach. The effect for women in overlap complaints is 8.6 percentage points (men is 15.3 pp), while the overall effect for women is 12.3 percentage points (men is 11.9 pp). This reveals a striking pattern: female induced reporters after #MeToo are positively selected across all possible group shares, whereas male induced reporters are *negatively* selected. In other words, men induced to report by #MeToo are more likely to make an unfounded accusation than are male always reporters. Moreover, for our preferred calibrations, the selection effect for men is so large that it dominates the treatment effect of adjudicators; men induced to report by #MeToo win less often than would male always reporters in a counterfactual world where #MeToo never happened.

Overall, our analysis suggests that adjudicators were more likely to believe sexual harassment claims by both men and women after #MeToo. This effect is particularly large for men. But the *newly induced* reporters exhibit a different pattern. Contrary to criticism that #MeToo encouraged women to report "bogus" cases, we find women started reporting sexual harassment complaints that were, on average, more likely to be founded. Men, on the other hand, started reporting sexual harassment complaints that were less likely to be founded.

1.1 Related Literature

There is a growing empirical literature studying the effects of the #MeToo movement on sexual wrongdoing. The closest study to ours is Levy and Mattsson (2025), which uses crime data to estimate the impact of #MeToo on sex crime reporting and arrests. They document a 10% increase in reported sex crimes and a corresponding rise in arrests for sexual assault, with larger effects for female victims. They interpret these findings as evidence of increased reporting, rather than increased incidence, of sexual violence. This conclusion is corroborated by Gauthier (2025), which finds that the #MeToo movement had the effect of increasing reporting while deterring sex crime. Our study builds on this work in two key ways. First, we focus on civil administrative cases, which capture harassment that is arguably more representative of the workplace incidents that #MeToo sought to address, rather than severe misconduct found in crime data. The milder nature of these complaints makes the question of substantiation particularly salient. While it is encouraging that #MeToo increased accountability for severe offenses, our data suggest an even larger response, perhaps reflecting a greater margin for change in the assessment of cases that we study.

Methodologically, we contribute a framework to decompose the ATT recovered from a difference-in-differences regression into a direct treatment effect and a selection effect, using two assumptions beyond the standard parallel trends assumption: an additional parallel trends assumption on non-selected individuals, and an assumption that treatment homogeneously affects all individuals whose outcome is determined after treatment. Thus, our framework is applicable to many policy settings that have a 'delayed-resolution' feature – public housing or grant applications, permit decisions, insurance claims, among others. While there is a rich literature dealing with treatment versus selection effects in IV settings (Kirkeboen et al., 2016; Mountjoy, 2022; Black et al., 2022) or RCTs (Lee, 2009), we are not aware of any analogous work for difference-in-differences.

Our claim that the choice to report is interdependent with the favorability of adjudicators is consistent with prior literature on sexual misconduct in the workplace, which finds that the rate of reporting is an equilibrium object. Boudreau et al. (2024) show that the costs of reporting sexual harassment, which are based in part on a fear of retaliation, are endogenous to the reporting and punishment institutions in use. For example, when outside job opportunities worsen, workers report less sexual harassment to avoid retaliation from the employer (Dahl and Knepper, 2026).

Women report more sexual harassment in male-dominated workplaces where high wages may be acting as a compensating differential (Folke and Rickne, 2022; Hersch, 2011, 2018), and they quit at higher rates following #MeToo (Batut et al., 2021). More broadly, the gender composition of managers within a firm affects the severity of consequences for cross-gender violence (not necessarily sexual) (Adams-Prassl et al., 2024), and therefore plausibly the rate at which wrongdoing occurs. In short, changing the costs or benefits of reporting may have large effects on the reporting rate. This paper studies what happens downstream once sexual harassment is reported and how changes in social attitudes impact who reports versus who is believed.

Our modeling approach also relates to the design of punishment mechanisms, as first studied by Becker (1968). Specifically, our theoretical framework borrows from Kaplow (2011); an adjudicator uses a threshold rule on the amount of evidence necessary for "believing" an accusation, where the threshold optimally trades off type I error (believing an unfounded allegation) and type II error (not believing a founded allegation). We make no attempt to explicitly model or measure the adjudicator's preferences over these types of errors. Instead, we allow #MeToo to flexibly change the evidentiary threshold through unmodeled changes to the costs of type I versus type II errors. An important implication is that our model allows for accusations to be "false,"[6] which differs from other models of the #MeToo movement (Cheng and Hsiaw, 2022; Boudreau et al., 2024).[7] Our model is also flexible enough to allow for changes in the cost of reporting sexual harassment after #MeToo, perhaps by reducing the risk of retaliation (Boudreau et al., 2024).

The effects we estimate reflect changes in individual behavior rather than changes in legal regimes. Numerous non-legal factors have been shown to affect judicial behavior, including lunch breaks, outcomes of recent sporting events, temperature, and economics training (Danziger et al., 2011; Eren and Mocan, 2018; Heyes and Saberian, 2019; Ash et al., 2026). But unlike these examples,

[6]Prior studies estimate that fewer than 10% of sexual assault complaints are false, in the sense that the alleged wrongdoing was not committed or attempted (Heenan and Murray, 2006; Lisak et al., 2010). In our data, 83% of cases that reach the investigation stage lose, suggesting far more complaints are rejected by adjudicators than would be expected based on the rate of false accusations alone. One factor that explains this discrepancy is that we study sexual *harassment*, not sexual *assault*. Second, not all unfounded complaints are false. Recalling the terminology used by the justice system, it is plausible that the typical complaint that loses in our setting is not false but rather baseless, i.e., it does not meet the legal definition of sexual harassment.

[7]By contrast, Chassang and Padró i Miquel (2018) theoretically study settings in which accusations can be unfounded. The focus of their investigation is on the optimal design of punishment policies when reporters may experience retaliation for bringing forth accusations. This relates to the consideration of adjudicators in the current paper, but our goal is to measure *how* adjudicators respond to #MeToo, not on describing the specific reasons for their responses.

the #MeToo movement was a society-wide, durable shift in attitudes towards sexual harassment that potentially affects adjudicators at scale – for example, female judges in China became more likely to grant divorces when domestic violence was indicated after #MeToo, corroborating that #MeToo can directly influence the behavior of judges by changing their attitudes (Cai et al., 2025).[8]

The remainder of the paper is organized as follows: Section 2 describes the #MeToo movement and institutional context, Section 3 describes our data, Section 4 presents our empirical strategy, Section 5 presents our aggregate estimates, Section 6 presents a model of the institutional setting, Section 7 empirically decomposes the combined effect into objects of interest, Section 8 presents the results of the decomposition, and Section 9 concludes. Proofs, additional tables and figures, and additional data details are housed in the Appendix.

2 Background

2.1 The #MeToo movement

Over half of women and around a fourth of men in the US report experiencing sexual harassment in their lifetime.[9] On October 15, 2017, the #MeToo movement gained widespread attention when actress Alyssa Milano tweeted asking people who had experienced sexual harassment to respond with the phrase "me too." Figure 1 shows that Google searches for #MeToo were roughly zero before October 2017 but spiked immediately thereafter.

The movement increased the salience of sexual harassment as a societal issue. #MeToo was used over 19 million times on Twitter in the first year alone (Anderson and Toor, 2018). In December 2017, Time magazine named #MeToo activists as their Person of the Year.

[8] In this regard, our paper also relates to recent research on the Black Lives Matter (BLM) movement, which studies changes in the number of reports, and subsequent rates, of crime (Campbell, 2024; Devi and Fryer Jr, 2020). However, to our knowledge, there has been no work relating BLM to the specific accusations of wrongdoing – that is, on whether accusations against specific police officers post-BLM are more or less likely to be founded, or on how adjudicators of police violence changed the way they assess accusations. Our model-based approach of disentangling the direct treatment and selection effects using "overlap" cases is likely applicable to these settings.

[9] In a 2018 Pew survey, 59% of women and 27% of men reported experiencing unwanted sexual advances or harassment (Pew Research Center, 2018). A 2019 poll revealed even higher numbers, with 81% of women and 42% of men reporting lifetime experiences of sexual harassment (Stop Street Harassment, 2019).

2.2 Legal protections against sexual harassment

Sexual harassment in employment is considered a form of sex-based discrimination under Title VII of the Civil Rights Act of 1964. The Equal Employment Opportunity Commission (EEOC) – the federal agency responsible for enforcing federal laws that protect against discrimination in employment – defines sexual harassment as "unwelcome sexual advances, requests for sexual favors, and other verbal or physical conduct of a sexual nature" that is made explicitly or implicitly a condition of employment, used as a basis of dismissal from employment, or interferes with the individual's work (Legal Information Institute).[10] State definitions of sexual harassment do not depart significantly from the EEOC definition; in fact, many states default to the EEOC definition. This homogeneity in definitions matters because it ensures that states classify similar incidents as harassment, allowing us to aggregate treatment effects meaningfully across states.[11]

2.3 Filing a sexual harassment complaint

Sexual harassment is typically treated as a civil offense and addressed through administrative complaints or civil lawsuits. However, serious instances of sexual harassment, such as rape, physical assault, or stalking, can rise to the level of a criminal offense or may be classified as a different type of sex crime. This paper focuses on civil complaints.

In civil cases, an individual ("complainant") can file a complaint against a respondent with an anti-discrimination agency at the state or federal level. We do not observe complaints filed with firms' Human Resources (HR) departments, nor do we observe whether an individual previously filed with these other outlets before filing with a state agency. According to officials who work at these state agencies, complainants may not report sexual harassment to their employer for one of the following reasons: they fear retaliation, they do not trust the HR department, a member of the HR department conducted the alleged discrimination, they want legal remedies for the situation, or their company does not have an HR department.

[10] Appendix B.1 provides specific examples of sexual harassment.

[11] Some states have expanded protections beyond federal regulations. These expansions include covering smaller employers, mandating specific training, allowing personal liability for harassers (whereas federal law only allows employers to be named in lawsuits), and explicitly protecting sexual orientation (Frisch, 2013).

3 Data

To obtain data on discrimination complaints, we filed Freedom of Information Act (FOIA) requests with all U.S. states that have relevant agencies (Arkansas and Mississippi do not). Our request sought all closed employment, public accommodation, housing, and education discrimination complaints filed between June 1, 2010, and June 1, 2023. We requested details including the respondent's name, filing and resolution dates, alleged basis and issue of discrimination, the agency's decision, and any compensation awarded. In Appendix B.1, we summarize the FOIA results by state. In Appendix E, we detail our FOIA request process.

Our main analysis focuses on employment discrimination. We harmonize the alleged bases of discrimination with federal discrimination categories: race, religion, sex (including sexual harassment), nationality, age, disability, and retaliation (punishment for filing a discrimination complaint).[12]

Our treatment group consists of sexual harassment complaints, and other types of discrimination form our control group. We drop a small number of sexual harassment cases not classified as sex discrimination, as well as sex discrimination that is not sexual harassment (e.g., pregnancy discrimination). Retaliation complaints that reference sex anywhere in the basis are treated as part of the treatment group, as they reflect backlash against sex-based allegations.

We exclude complaints lasting over 764 days (more than two standard deviations above the mean of 262 days). Our final sample contains 191,985 employment discrimination complaints. We use complaints from public accommodation, housing, and education sectors as a robustness check, and this larger sample contains 345,000 complaints.

After filing, a discrimination complaint can be closed in multiple ways. A complaint may be dismissed, be adjudicated in court, be settled by the parties (where no determination of liability is rendered but the complainant can receive compensation), or proceed to investigation to determine whether or not there was probable cause that discrimination occurred. The investigation is conducted by agency officials and may include interviews with relevant parties, evidence gathering, or

[12]Cases whose bases do not fit into the main categories, such as genetic, police, veteran, drug, and other types of miscellaneous discrimination comprise 2% of the data and are excluded from the analysis. For discrimination complaints that have multiple bases, we take the first basis if the state does not provide priority ordering. https://www.eeoc.gov/employers/small-business/3-who-protected-employment-discrimination

site visits. Figure 2 presents a flowchart of the complaint process.[13]

Our main outcome of interest is the finding of the investigation stage, as the investigation finding offers a clear signal as to the level of evidence supporting the complainant. Therefore, we restrict our primary analysis to cases that reach the investigation stage. We consider a complainant to have "won" if there is a finding of reasonable or probable cause of discrimination during the investigation. We consider the complainant to have "lost" if no probable cause is found. If the complainant wins, potential remedies include the defendant paying compensation to the victim, the employer implementing new policies, or other arrangements to compensate the complainant. Since these are civil proceedings, no criminal charges or jail time are involved. Data on remedies was sparsely provided by states, and we therefore do not consider it as an outcome in our analysis.

We also study other outcomes in the Appendix, which are qualitatively different because they do not terminate following an investigation at the agency level. We consider a case "Settled" when the raw data indicates a settlement, mediation, or withdrawal with adjustment. A case ends in "Court" if the case is removed to a court proceeding overseen by a judge or magistrate. We do not observe the outcomes of complaints that were removed to court. We define a case as "Dismissed" if it is administratively closed, withdrawn (without settlement), or labeled as "Dismissed" in the raw data. Administrative closures include cases with logistical issues, such as being out of the agency's jurisdiction, being filed in an untimely manner, or being unable to contact the complainant. No restitution is made in such cases, as they exit prior to a thorough investigation.[14] These outcomes are mutually exclusive, with a few rare exceptions that arise from exigencies of each state's system. We provide additional details on mapping the raw data into case outcomes in Appendix E.

Approximately 7% of cases settle, 8% are dismissed outright, and 43% go to court, where we do not observe their outcome. Of the remaining 42% of cases that are investigated by the agency, 83% (0.35 / 0.42) lose. Table 1 provides summary statistics and differences in means before and after

[13]The outcomes in our data are terminal, meaning we observe the discrimination ruling for every investigated case. Moreover, outcomes are non-duplicative for all but 0.17% of cases, meaning that nearly every investigated case receives only one outcome code in our data. Our results are robust to excluding the small number of cases with multiple outcomes.

[14]Complaints are often dismissed for administrative reasons, including when the complainant cannot be located despite assiduous efforts, when the complainant or defendant has declared bankruptcy or passed away, when the complainant has filed an equivalent federal complaint, or when the complainant refuses to cooperate in the resolution process. Some complaints are dismissed for insufficient evidence, but the evidentiary threshold for doing so is likely very low given the small percentage of cases that are dismissed outright.

#MeToo. In Appendix B.1, we detail trends in the rate of sexual harassment allegations over time, both before and after #MeToo.

4 Empirical strategy

We estimate the effect of #MeToo on sexual harassment complaints using difference-in-differences. Treatment timing is the start of the #MeToo movement on October 15, 2017 and is constant across all states. Our main estimating equation is a two-way fixed effects specification within a state:

$$Y_{is(i)m(i)} = \beta^{CE}\left(\mathbb{1}\left[g(i) = \text{SH}\right] \times \mathbb{1}\left[m(i) \geqslant \text{Post}\right]\right) + \alpha_{g(i)s(i)} + \mu_{m(i)s(i)} + \epsilon_{is(i)m(i)} \tag{1a}$$

where i is a complaint, $g(i) \in \{\text{sexual harassment, age, disability, nationality, race, religion, retaliation}\}$ indexes complaint type, $s(i)$ indexes state, and $m(i)$ indexes the year-month of when the complaint was resolved.[15] Time fixed effects $m(i)$ are defined at complaint resolution rather than complaint filing to capture changes in adjudicator behavior at the time of decision. Post denotes October 15, 2017 when #MeToo began. The treatment indicator $\mathbb{1}\left[g(i) = \text{SH}\right] \times \mathbb{1}\left[m(i) \geqslant \text{Post}\right] = 1$ if i is a sexual harassment complaint resolved after #MeToo. $Y_{is(i)m(i)}$ is an indicator for whether the complainant wins their case (versus loses), conditional on reaching the investigation stage. This regression therefore compares the win rate of sexual harassment complaints (treated) to the win rates of other types of discrimination complaints (control) in the same state, before and after #MeToo. Following Levy and Mattsson (2025), we cluster standard errors at the state by complaint type level to align with how treatment is assigned, and to account for within-state correlation in complaint adjudication, for a total of 57 clusters.

We include interactions with state dummies $s(i)$ to account for differences in how complaints are adjudicated between states. $\alpha_{g(i)s(i)}$ are complaint type-by-state fixed effects accounting for constant state policies that differentially affect sexual harassment and non-sexual harassment outcomes. $\mu_{m(i)s(i)}$ are time-by-state fixed effects accounting for time-varying state policies that affect all cases the same. For example, in the six years following the start of #MeToo, half of all states enacted additional anti-harassment laws to further protect workers (National Women's Law Center, 2023).

[15] Since some retaliation cases may be filed in response to sexual harassment, in Appendix C, we drop retaliation complaints from the control group and confirm this does not change our results.

Since these laws applied to all discrimination cases, $\mu_{m(i)s(i)}$ control for the impact of these policies.

β^{CE} ("combined effect") is our main coefficient. In order to interpret this coefficient as a parameter of interest, we assume throughout that #MeToo is unanticipated (no effect in the pre-periods), and that in the absence of #MeToo-induced changes in case adjudication, the outcomes of sexual harassment and control complaints would have trended in parallel within a state. To formalize, assume a simple two period model where $t = -1$ is the pre-period and $t = 0$ is the post-period. Suppressing state indices, let $Y_{it}(1)$ be the potential outcome if a complaint had received treatment in time t and $Y_{it}(0)$ if a complaint had not received treatment in time t. Let D_{it} indicate actual treatment status at time t, so $D_{i0} = 1$ for (and only for) sexual harassment complaints resolved after #MeToo. Under a *standard parallel trends assumption* that $\mathbb{E}[Y_{i0}(0) - Y_{i-1}(0)|D_{i0} = 0] = \mathbb{E}[Y_{i0}(0) - Y_{i-1}(0)|D_{i0} = 1]$, β^{CE} identifies the average treatment effect on the treated, $ATT^{CE} := \mathbb{E}[Y_{i0}(1) - Y_{i-1}(0)|D_{i0} = 1] - \mathbb{E}[Y_{i0}(0) - Y_{i-1}(0)|D_{i0} = 1]$.

We call β^{CE} the combined effect because it captures the effect of changes in the probability of winning holding case characteristics fixed (direct treatment), as well as the effect of changes in the composition of cases reported (selection).[16] The sign of β^{CE} is, by itself, insufficient to determine the direction or magnitude of these two effects. To see why this is the case, suppose #MeToo makes adjudicators more likely to believe all sexual harassment complainants. This implies a positive direct treatment effect. But we may still recover a *negative* β^{CE} if #MeToo induces mass reporting of grievances that do not rise to the level of sexual harassment. In other words, a sufficiently high share of new unfounded complaints may result in a negative β^{CE}, even if adjudicators' attitudes overall are more favorable to complainants. Analogously, a sufficiently high share of new founded complaints may result in a positive β^{CE}, even if adjudicators' attitudes overall are less favorable to complainants. Decomposing β^{CE} will be the task of the second half of this paper.

For our event studies, we estimate a dynamic version of Equation 1a, where t indexes yearly leads and lags around #MeToo, $t = -1$ is the 12 month period before October 15, 2017, and SH is a

[16]These changes may have occurred for several reasons, including as a result of state policies instituted after #MeToo that may have *differentially* impacted sexual harassment and control complaints. Because our fixed effects would not absorb such differential impacts, they would appropriately be captured in β^{CE} as part of the #MeToo effect.

sexual harassment complaint:

$$Y_{is(i)m(i)} = \sum_{t \neq -1} \beta(t) \left(\mathbb{1}\left[g(i) = \text{SH}\right] \times \mathbb{1}\left[m(i) \in t\right] \right) + \alpha_{g(i)s(i)} + \mu_{m(i)s(i)} + \epsilon_{is(i)m(i)} \tag{1b}$$

In the third specification, we estimate the differential impact of #MeToo for women versus men using triple differences. Specifically, we estimate a version of Equation 1a but interacted with whether the complainant is female:

$$\begin{aligned} Y_{is(i)m(i)} = \; & \beta^{DDD} \left(\mathbb{1}\left[g(i) = \text{SH}\right] \times \mathbb{1}\left[m(i) \geqslant \text{Post}\right] \times \mathbb{1}\left[i = \text{Female}\right] \right) \\ & + \beta_M^{CE} \left(\mathbb{1}\left[g(i) = \text{SH}\right] \times \mathbb{1}\left[m(i) \geqslant \text{Post}\right] \right) \\ & + \alpha_{g(i)s(i)} \; + \; \mu_{m(i)s(i)} \\ & + \left(\alpha_{g(i)s(i)} \times \mathbb{1}\left[i = \text{Female}\right] \right) \; + \; \left(\mu_{m(i)s(i)} \times \mathbb{1}\left[i = \text{Female}\right] \right) + \epsilon_{is(i)m(i)} \end{aligned} \tag{2a}$$

We assume that the outcomes of sexual harassment and control cases brought by *male* complainants trended in parallel within a state, under which β_M^{CE} identifies the ATT for male complainants. To identify $\beta_W^{CE} := \beta^{DDD} + \beta_M^{CE}$ as the ATT for female complainants, we make an analog of the parallel trends assumption: that in the absence of #MeToo, the difference in outcomes for sexual harassment versus control complaints filed by women would have trended in parallel to the difference in outcomes for sexual harassment versus control complaints filed by men (Olden and Møen, 2022).

Finally, we estimate the dynamic version of Equation 2a:

$$\begin{aligned} Y_{is(i)m(i)} = \; & \sum_{t \neq -1} \beta^{DDD}(t) \left(\mathbb{1}\left[g(i) = \text{SH}\right] \times \mathbb{1}\left[m(i) \in t\right] \times \mathbb{1}\left[i = \text{Female}\right] \right) \\ & + \sum_{t \neq -1} \beta_M^{CE}(t) \left(\mathbb{1}\left[g(i) = \text{SH}\right] \times \mathbb{1}\left[m(i) \in t\right] \right) \\ & + \alpha_{g(i)s(i)} + \mu_{m(i)s(i)} \\ & + \left(\alpha_{g(i)s(i)} \times \mathbb{1}\left[i = \text{Female}\right] \right) \; + \; \left(\mu_{m(i)s(i)} \times \mathbb{1}\left[i = \text{Female}\right] \right) + \epsilon_{is(i)m(i)} \end{aligned} \tag{2b}$$

5 Aggregate Results

We present estimates of Equation 1a in Table 2, Column 1, and the event study that corresponds to Equation 1b in Figure 3a. Our FOIA request covered the 7 years prior and the 6 years after #MeToo, so the x-axis of each event study spans from -7 to 6 (more details on the FOIA process are in Appendix E). We estimate that #MeToo raised complainant win probability by $\hat{\beta}^{CE} = 10.1$ percentage points (std. err. = 2.9) or 66% over the control mean. A visual inspection of the event study shows that there is little evidence of pre-trends and that pre-treatment coefficients are not statistically significant. Post-treatment coefficients are positive and indicate an average impact of #MeToo that aligns with the point estimate of the static regression.

In Appendix C, we show that our findings are robust to multiple changes in specifications and data choices, including: restricting to complaints alleging only one type of discrimination to eliminate potential ambiguity from overlapping discrimination types; excluding retaliation complaints from the control group in case they are responses to previous sexual harassment; restricting to complaints filed pre-Covid-19 to ensure our main effect is not being driven by pandemic-related disruptions and to remove mechanical over-representation of quick-to-resolve complaints at the end of our sample; re-running our main results with time fixed effects defined at case filing rather than case resolution; and extending our main analysis to a broader set of discrimination complaints (including housing and education) to test whether our main effect can be replicated outside of employment. $\hat{\beta}^{CE}$ remains positive and significant in every robustness check, retaining a similar order of magnitude as our preferred specification. Finally, we conduct a permutation test where we reassign treatment to untreated discrimination types.

We show results and event studies for secondary outcomes—rates of settling and exiting to court—in Appendix B.2. #MeToo had virtually no impact on the propensity to settle and a minor increase in the likelihood of exiting to court (0.7 percentage points, std. err. = 0.6) that is unlikely to be economically significant.[17]

[17] More specifically, we claim that #MeToo's effect on the selection of cases into exiting to court is unlikely to significantly drive the effect of #MeToo on win rates. This can be seen via a Lee (2009)-style bounding exercise on the excess court cases post #MeToo: we reassign the 0.7pp excess court cases post #MeToo to either all be wins or all be losses to mimic the extremal counterfactuals in which excess court cases instead went to investigation. Under this exercise, the estimate of our main effect, which we describe shortly, of 0.101 would change to 0.100 (excess court cases would all have been losses) and 0.102 (excess court cases would all have been wins), suggesting that changes to the likelihood of exiting to court only minimally affects our results.

Next, we examine heterogeneity by complainant gender. Since we only observe complainant gender in a subset of states, we first assess whether these states differ systematically from those that did not provide gender data. In Column 2 of Table 2, we re-estimate Equation 1a only in the subset of states with gender information. The estimated effects are similar to those in the full sample: the effect on winning is nearly identical (0.101 vs. 0.116). This similarity provides reassurance that our main results are not driven by differences in sample composition across states.[18]

A priori, we expected any effect of #MeToo to be largest for female complainants. This is not the case. In Column 3 of Table 2, we show the results of estimating Equation 2a. We find that female complainants do not benefit more from #MeToo compared to male complainants. #MeToo increases the probability that a man wins his case by $\hat{\beta}_M^{CE} = 0.119$ (std. err. = 0.026), which is not significantly different from women ($\hat{\beta}_W^{CE} - \hat{\beta}_M^{CE} = 0.004$, std. err. = 0.022). Figure 3b presents the corresponding event study, again exhibiting little evidence of pre-trends. The orange coefficients for women closely align with the blue coefficients for men, and the average of the post-period coefficients roughly equals the point estimate. The only exception is at $t = 6$, but we interpret this as an artifact of few observations with gender in the last period rather than as economically meaningful (note the larger standard errors on the male coefficient at $t = 6$). More broadly, in Appendix B.2, we show that no meaningful gender differences exist across any outcomes.

Finally, we examine heterogeneity by state in Appendix B.3, grouping states into red, blue, or purple by their voting history (see Table A1 for the classification). We find that our results are driven by blue (i.e., Democrat-voting) states. The effect of #MeToo in red and purple states is virtually zero. Our finding aligns with prior literature on the geography of #MeToo, which found that the positive effects of #MeToo for victims is concentrated in liberal places in the United States (Levy and Mattsson, 2025).

[18]States with and without gender information also do not notably differ in their female labor force participation (FLFP) rates or gender wage gaps. States that provided gender information had an average FLFP and gender wage gap of 58% and 86%, respectively, compared to 57% and 83% for states that did not provide gender information (U.S. Bureau of Labor Statistics, 2023, 2024).

6 Theoretical Framework

Our goal in this section is to decompose the combined effect of #MeToo into a selection effect — how the foundedness of complaints changes after #MeToo — and an adjudicator treatment effect — how #MeToo affects adjudicator favorability toward complainants. To do so, we begin by presenting a simple theoretical framework. In it, each individual considers whether to make a report of sexual harassment with the pre-existing knowledge of whether their potential accusation is founded or unfounded, where a founded accusation is one that both accurately captures events experienced by the complainant and rises to the legal definition of sexual harassment. Making an accusation is costly but leads to probabilistic payments if their accusation is believed by an adjudicator. We use the model to study the optimal decision to report instances of (potential) sexual harassment and to understand adjudicators' decisions to "believe" the accusations or not. We then consider how changes to parameters of the model after #MeToo affect these decisions, which allows for a decomposition result that we can take to the data.

After presenting the simple model, we discuss how its predictions are robust to a number of alternative assumptions. For example, while we present a model where individuals know whether their potential accusation is founded or not before they decide to report, our findings are essentially unchanged if instead individuals are unsure of the foundedness of their claim, i.e., if they do not know whether an experience rises to the level of sexual harassment. The robustness of our theoretical findings to realistic alternatives suggests that our empirical decomposition is likely to faithfully capture the effects of #MeToo on adjudicators and on selection into reporting.

6.1 Model Setup

There are two time periods $t \in \{-1, 0\}$, where $t = -1$ corresponds to the time period before #MeToo and $t = 0$ corresponds to the time period after #MeToo. In each period, there is a unit mass of individuals forming set I^t whom we think of as the set of people who could plausibly report sexual harassment. Each individual $i \in I^t$ is active only in time period t and has a type $\theta_i = (c_i, \tau_i) \in \{L, H\} \times \{F, U\}$, where $0 = L < H < \infty$ is the cost of making an accusation, and F and U stand for whether a potential accusation would be "founded" or "unfounded," respectively. We assume that in each time period, there exists a positive measure of agents of each

type $\theta \in \{L, H\} \times \{F, U\}$.

The timing of the stage game is as follows. Upon observing her type, each individual chooses to either *report* or *not report*. If an individual does not report, then her game ends immediately and she receives utility normalized to zero. If she reports, her case enters the review process. Conditional on reporting, one of three mutually exclusive outcomes occurs to the case: (i) it is settled with probability p^t_{settle}, (ii) it is routed to court with probability p^t_{court}, or (iii) it proceeds to the agency investigation stage with the remaining probability $p^t_{inv} := 1 - p^t_{settle} - p^t_{court}$. We assume p^t_{settle} and p^t_{court} (and hence p^t_{inv}) are independent of τ_i and other case features, though expected payments may depend on τ_i. Moreover, we assume that $p^t_{inv} > 0$ for all t.

For cases that enter the investigation stage, two mutually exclusive outcomes occur. The investigation can either be "won" by the complainant, in which case she receives a monetary award π^t_{win}, or it can be "lost" by the complainant, in which case no payment is made. The award π^t_{win} is potentially random, and the distribution of awards can differ by the foundedness of the claim, so long as the award payment has finite expectation for each level of foundedness. "Winning" is a stochastic outcome that occurs with probability $p^t_{win}(\tau_i)$ in period t.

To determine whether a case is won or lost, the investigator relies on the level of (testimonial) *evidence* $e_i \in [0, 1]$ generated during the investigation in favor of complainant i. As evidence is generated during the investigation, its realization is not observed in prior stages of the game, notably, at the time the individual decides to make a report or not. A higher level of evidence implies the complainant's case is more likely to be founded; each complaint generates a level of evidence that is randomly and independently distributed, where the distribution depends on the foundedness of the complaint. To formalize these notions regarding evidence, let G_F and G_U be distributions of evidence generated for founded and unfounded claims, respectively. We assume that G_F and G_U have full support over $[0, 1]$ and are continuously differentiable with probability distribution functions g_F and g_U, respectively. We assume that the likelihood ratio $g_F(e)/g_U(e)$ is strictly increasing in e, that is, higher values of evidence are more likely to be uncovered if the complaint is founded, i.e., more evidence supports the complaint. Following Kaplow (2011), we assume that there is a (time-varying) *evidentiary threshold* $\underline{e}^t \in [0, 1]$ such that a case is won in time period t if and only if the complainant generates at least $\underline{e}^t$ level of evidence. Given that

reaching the investigation stage is independent of the case foundedness, the probability that case i wins conditional on reaching the investigation stage is $1 - G_{\tau_i}(\underline{e}^t)$. Because the level of evidence generated is a noisy signal of the foundedness of a complaint, for any $\underline{e}^t \in (0,1)$, there is positive probability of both type I error (unfounded complaints that are won, which occurs for unfounded complaints with probability $1 - G_U(\underline{e}^t)$) and type II error (founded complaints that are lost, which occurs for founded complaints with probability $G_F(\underline{e}^t)$).[19]

Each individual's payoff from reporting is equal to any monetary award she receives minus the cost of reporting, and her payoff from not reporting is normalized to zero.

Let

$$u_i^t := p_{settle}^t \, \mathbb{E}_t[\pi_{settle}^t | \tau_i] + p_{court}^t \, \mathbb{E}_t[\pi_{court}^t | \tau_i] + p_{inv}^t \, p_{win}^t(\tau_i) \, \mathbb{E}_t[\pi_{win}^t | \tau_i] - c_i \tag{3}$$

represent individual i's expected payoff from reporting. An individual i in time t files a report if and only if $u_i^t \geqslant 0$. Note that if an individual i files a report at time t, then so do all the other individuals of the same type at time t. Therefore, we abuse notation and refer to the expected utility of reporting of all agents of type θ in time t as u_θ^t. Note that in our empirical analysis, our main outcome corresponds to $p_{win}^t(\tau)$, i.e., the probability of a favorable finding conditional on reaching the investigation stage, rather than the unconditional probability $p_{inv}^t \cdot p_{win}^t(\tau)$.

6.2 Discussion of Modeling Assumptions

Our stylized model is meant to capture the key intuitions and mechanisms necessary to tie back to our empirical setting. Our goal, therefore, is not to maximize generality, but rather readability. In this section, we discuss some modeling assumptions and whether they affect key takeaways.

First, our model considers only two time periods, one before #MeToo and one after. Our empirical specification considers multiple times in each of the pre- and post-#MeToo periods, and our theoretical results would remain similar but be more notationally complex with more periods.

[19] Kaplow's model takes as primitives (potentially heterogeneous) costs on adjudicators for type I vs type II errors. Because of the monotone likelihood ratio assumption on the evidence generating process, any collection of costs of type I and type II errors induces a threshold rule as being optimal. We do not explicitly model these error costs, and instead assume that #MeToo can lower (raise) the evidentiary threshold used, which can occur by either increasing (decreasing) the cost of type II error or decreasing (increasing) the cost of type I error. In what follows, we aim to empirically identify #MeToo's effect on the threshold.

Second, our model assumes four types of potential reporters, as indexed by cost and "foundedness." While we find it reasonable, based on the literature, to dichotomize the foundedness of claims, there is no reason to limit the support of costs. What is important for our analysis is to ensure that there are positive measures of both founded and unfounded reports in each period. Our assumptions that low-cost types face a cost $L = 0$ and that there exist positive measure sets of low-cost founded and unfounded cases ensure the desired outcome. Moreover, our assumption that individuals observe their type fully prior to the decision to report implies that accusers know ex-ante whether or not their complaint is in good faith. This is not necessary, and we could have instead complicated our model by assuming individuals receive a (potentially uninformative) signal of the foundedness of their potential allegation. This alternative would be more realistic if individuals are not sure whether a potential interaction rises to the level of sexual harassment in the eyes of the law when deciding to report. As with the restriction on the number of types, this alternative does not affect our takeaways because our theoretical results take as given the set of reporters pre- and post-#MeToo. That is, our model provides an economic justification for why an individual would or would not report in different time periods, but our results do not rely on the precise details of why individuals choose to report or not.

Third, and relatedly, we make no attempt to model the decision of a wrongdoer to harass. Plausibly, #MeToo had a deterring effect on wrongdoing. Our model encodes whether a particular individual was or was not harassed into their type via the second component $\tau \in \{F, U\}$. Therefore, deterrence (or any other forces changing the amount of wrongdoing) is represented as a change in the distribution of types between the time periods. Note that we can also allow the precise costs paid by each type to vary across periods without affecting our takeaways.

Fourth, we assume that the probabilities of going to court or settling are independent of other features of the model. This implies that cases that go to investigation are neither positively nor negatively selected for foundedness.[20] While this is a limiting assumption, it is common in the literature[21] and our theoretical results are not knife-edge—that is, they remain qualitatively valid if

[20] As previously discussed, complainants of dismissed cases receive no restitution. As dismissal typically occurs for administrative reasons, and occurs before the investigation stage, we exclude this outcome from our model.

[21] Specifically, it is common in the literature to assume that previous actions in multi-stage assessment systems do not affect the analysis of the analyzed stage. For example, Arnold et al. (2018) study the bail decisions of judges and whether they exhibit racial preferences. However, the authors' estimates assumes that judges' choices do not reflect the actions of previous actors who interact with the defendant. As they state on page 1902, a "concern is that bail judges may be

there is a sufficiently small (positive or negative) correlation between going to court or settlement and a complaint's foundedness. Reassuringly, as discussed in Footnote 17, we do not view the selection of filed cases into the investigation stage to be first order.

Finally, we do not include any notion of demographic groups in our model, while our empirical analysis separately considers the differential effects of #MeToo on male versus female complainants. Our model can straightforwardly be extended to consider different demographic groups (e.g. men and women), and all of our findings below will hold within group. Moreover, the types and type distributions in each of $t = -1$ and $t = 0$ can be different across groups. This flexibility allows us to consider the case in which, for example, the type distributions of male and female reporters differ pre-#MeToo, and where #MeToo differentially affects the type distribution—by deterring acts of sexual wrongdoing, or changing the cost of reporting—for each gender.

6.3 Theoretical Tests

Recall that our key aim with the model is to decompose the total effect of #MeToo on the complainant win rate into a selection into reporting effect and a treatment of adjudicators effect. To aid us in doing so, we introduce some notation and terminology.

For each period $t \in \{-1, 0\}$ and each $\tau \in \{F, U\}$, let R^t_τ represent the set of reporters with foundedness level τ. Because of the assumption of finitely many types and full support, each R^t_τ is (Lebesgue) measurable and has positive (Lebesgue) measure, which we denote by $|\cdot|$, i.e., $|R^t_\tau| > 0$. Therefore,

$$\mathcal{T}_t := \frac{|R^t_F|}{|R^t_F| + |R^t_U|} \in (0, 1)$$

fraction of complaints in period t are founded, and $R^t_F, R^t_U > 0$ (i.e., that both founded and unfounded accusations are made in each period) ensures that $\mathcal{T}_t$ is strictly between 0 and 1.

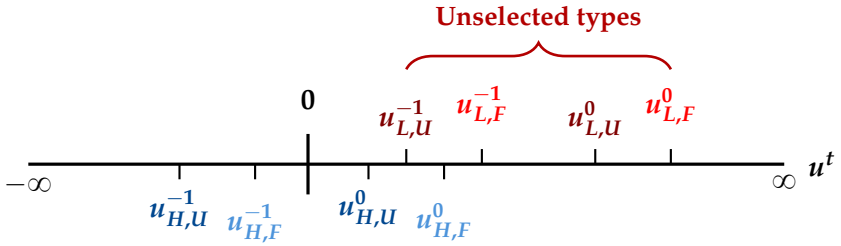
Let $r^0 \subset R^0_F \bigcup R^0_U$ be a positive measure set of reporters post-#MeToo, where we represent the subset of founded and unfounded reports within set r^0 as r^0_F and r^0_U, respectively. We call set r^0

influenced by other court actors (e.g., prosecutors) when making decisions..." Our assumption similarly attributes the effects of changes in the win rate after #MeToo to adjudicators, however, as discussed in Section 6.2, our framework allows #MeToo to change the foundedness distribution, potentially through deterrence of wrongdoing.

unselected if

$$\frac{|r_F^0|}{|r_F^0| + |r_U^0|} = \mathcal{T}_{-1}.$$

In words, a set of reporters r^0 in the post-#MeToo period is said to be unselected if the share of them with founded accusations is equal to the share of founded accusations in the pre-#MeToo period. To gain intuition into the properties of such a set, consider an example in which the distribution of types is the same in both periods, i.e., that #MeToo did not affect reporting costs or affect the incidence of sexual harassment. Instead, suppose that #MeToo raised the expected utility from reporting for each type, $u_\theta^0 \geqslant u_\theta^{-1}$ for all types θ. As depicted graphically below, this means that all types who reported pre-#MeToo continue to report after #MeToo (e.g. types (L, U) and (L, F) in dark and light red, respectively), and some types that did not report pre-#MeToo switch to doing so after #MeToo (e.g. types (H, U) and (H, F) in dark and light blue, respectively). The collection of types that report in both periods form an unselected subset of reporters post-#MeToo, while the new types who report can be either positively selected (i.e. more likely to be founded) or negatively selected (i.e. less likely to be founded).



Departing from this simple example, our definition of an unselected subset of reporters holds regardless of changes in the composition of types, because we place no restriction on the measure of the set r^0. In other words, an unselected subset of reporters always exists. Call an unselected r^0 *always reporters (ARs)* if it is a maximal unselected set, i.e., there does not exist another unselected set of reporters $\tilde{r}^0$ such that $|\tilde{r}^0| > |r^0|$. Given a set r^0 of always reporters, we refer to the remaining reporters post-#MeToo $R^0 \backslash r^0$ as *induced reporters (IRs)*. Because there is a finite measure of reporters in each period, a set of always reporters, and therefore induced reporters, always exists. While there are, in general, many possible sets of always reporters, they each will have the same composition of foundedness as the entire set of reporting agents pre-#MeToo, and our upcoming results will

rely only on the composition of the set of always reporters, not their identities. If there are no compositional changes across periods and the same types report in both periods, all reporters post-#MeToo are always reporters.

The following proposition relates changes in the win rate of different groups to the selection of individuals into reporting and to changes to the adjudicators' evidentiary threshold, and its proof is contained in Appendix A.

Proposition 1.

1. *If the expected win rate increases (decreases) for always reporters after #MeToo, then the evidentiary threshold used by adjudicators is lowered (raised) after #MeToo.*
2. *Induced reporters are more likely to be founded than are always reporters if and only if induced reporters have higher expected win rates than always reporters after #MeToo.*

Part 1 discusses how the impact of the win rate on always reporters reveals changes to the evidentiary threshold: an increase in win rates for always reporters corresponds to a lowering of the evidentiary threshold. Similarly, a decrease in win rates for always reporters corresponds to a raising of the evidentiary threshold. Part 2 discusses how selection into reporting affects outcomes of interest. If induced reporters are positively selected (their complaints are more likely to be founded), then we expect win rates to be higher for induced reporters after #MeToo than for always reporters. If induced reporters are negatively selected, then we expect the opposite.

Remark 1. *Proposition 1 is stated in terms of always reporters and induced reporters. However, its findings and the arguments behind them are unchanged if we replace these two groups with any unselected subset of reporters and the complement of that unselected subset, respectively. We state the result in terms of always and induced reporters to better comport with our upcoming decomposition.*

7 Empirical Decomposition

We now turn to empirically decomposing our main difference-in-differences effect into a direct treatment effect on adjudicators, and into a selection-into-reporting effect. This decomposition is based on our theoretical framework, which partitions reporters post-#MeToo into a set of always reporters and induced reporters. In this section, we demonstrate an empirical method for estimating

effects on each of these groups. Then, in Section 8, we use these estimates, along with the findings of Proposition 1, to make conclusions about selection into reporting and changes to adjudicators' evidentiary threshold.

Again considering a two-period model for simplicity of exposition, recall that the "combined" ATT is given by $ATT^{CE} = \mathbb{E}[Y_{i0}(1) - Y_{i-1}(0)|D_{i0} = 1] - \mathbb{E}[Y_{i0}(0) - Y_{i-1}(0)|D_{i0} = 1]$. This is the combined effect of #MeToo on all reporters. To aid in our decomposition, define the ATT for always reporters and induced reporters, respectively, as $ATT^{AR} := \mathbb{E}[Y_{i0}(1) - Y_{i-1}(0)|D_{i0} = 1, ARs] - \mathbb{E}[Y_{i0}(0) - Y_{i-1}(0)|D_{i0} = 1, ARs]$ and $ATT^{IR} := \mathbb{E}[Y_{i0}(1) - Y_{i-1}(0)|D_{i0} = 1, IRs] - \mathbb{E}[Y_{i0}(0) - Y_{i-1}(0)|D_{i0} = 1, IRs]$. Each of these two terms is similar to ATT^{CE} conditioning on always/induced reporters, respectively. Let $\omega \in (0, 1)$ be the share of always reporters in the post period.

The Law of Iterated Expectations yields the following decomposition of the combined effect of #MeToo, ATT^{CE}, into the causal effects of #MeToo on ARs and IRs:

$$ATT^{CE} = \omega ATT^{AR} + (1 - \omega) ATT^{IR}. \tag{4}$$

As a reminder, our setting with pooled cross-sectional data means that the definition of always (and induced) reporters depends on the composition of case foundedness: always reporters are neither positively nor negatively selected (compared to the entire set of reporters pre-#MeToo). Because the foundedness of any given case is not directly observable, the share of always reporters (ω) is a latent variable. As a result, novel identification and estimation exercises are required to discern the key terms of interest in Equation 4:

- In Section 4 we have shown that ATT^{CE} is identified by β^{CE} (see Equation 1a) under the *standard parallel trends assumption*, and presented empirical estimates.
- In Section 7.1, we show that ATT^{AR} is identified (under plausible assumptions)[22] by restricting our regression analysis to the subset of cases filed before, but resolved after, #MeToo.
- In Section 7.2, we present calibrations of ω based on a bounding exercise, and a relevant estimate from the literature.

[22] Because of our difference-in-differences empirical strategy, the primary assumption required is a refinement of the standard parallel trends assumption.

- With these terms in hand, we identify ATT^{IR} as the unique solution to Equation 4.

Section 8 relates these individual estimates to our economic questions of interest: do adjudicators believe complainants more after #MeToo, and are the complaints of induced reporters more founded on average than the complaints of always reporters?

Note that we have not formally discussed analogous gender-specific decompositions. However, Equation 4 is valid for decomposing β_M^{CE} and β_W^{CE}, appropriately indexing the ATTs and ω with gender-specific terms.

7.1 Estimating ATT of always reporters

In this section, we show how to empirically estimate ATT^{AR} by exploiting a feature of discrimination complaints shared with many other settings (such as permit approvals or grant applications); namely, that there is a delay between filing and resolution.

Specifically, we consider complaints filed *before* but resolved *after* #MeToo. We call these *overlap* cases. Since overlap cases are filed before #MeToo, their selection into reporting could not have been affected by #MeToo. Under plausible assumptions, we show that the treatment effect for overlap complaints identifies the treatment effect for always reporters, i.e., it reflects the change in leniency on the side of the institutional adjudication process. We often refer to this effect as the direct treatment effect on adjudicators, while recognizing that other institutional changes affecting leniency of the investigation may contribute to this direct effect (e.g., more favorable treatment of complainants by staffers or witnesses).

To operationalize the overlap analysis, we run the same difference-in-differences specifications as our main regressions (Equations 1a and 1b), but restrict the sample to cases filed before #MeToo. Therefore, this restricted regression compares overlap and pre-period sexual harassment cases to overlap and pre-period control cases, where we refer to the coefficient of interest as β^O (where the superscript denotes overlap cases). Figure 4 illustrates this comparison.

To identify β^O as ATT^{AR}, we make two assumptions, A1 and A2, which we describe below:

A1 Parallel Trends for ARs: $\mathbb{E}[Y_{i0}(0) - Y_{i-1}(0)|D_{i0} = 1, ARs] = \mathbb{E}[Y_{i0}(0) - Y_{i-1}(0)|D_{i0} = 0]$, or that the untreated potential outcomes of sexual harassment always reporters trend the same

as the untreated potential outcomes of control group reporters.

Note that A1 is equivalent to the *standard parallel trends assumption* presented in Section 4, except that it applies only to the outcomes of always reporters in the treated group as opposed to all reporters in the treated group. Intuitively, this difference in assumptions mirrors the direct treatment and selection effects we study. The standard parallel trends assumption, along with ATT^{CE}, allow for selection into reporting; the standard parallel trends assumption assumes that in the absence of the direct treatment effect of adjudicators, the expected win rate of *the entire set of sexual harassment cases post #MeToo* trends the same as the win rate of non-sexual harassment cases. Assumption $A1$, along with ATT^{AR}, do not allow for selection into reporting; because always reporters are definitionally unselected, A1 states that in the absence of the direct and selection effects, the expected win rate of sexual harassment cases trends as does the win rate of non-sexual harassment cases. Since A1 is a parallel trends assumption, plotting and examining pre-trends can provide evidence consistent or inconsistent with this assumption. The absence of pre-trends, which we verify below, is consistent with A1.

Recalling terminology from our theoretical model, overlap cases are definitionally unselected, but practically, they account for only a subset of always reporters. This is because there are many always reporter cases filed after #MeToo. The following assumption is necessary to extrapolate the treatment effect from overlap cases to the entire set of always reporters, regardless of filing timing:

A2 Filing Invariance: Treatment effects for any two always reporter cases resolved after #MeToo are the same. This implies that cases filed before #MeToo but resolved after have the same treatment effects as always reporter cases both filed and resolved after #MeToo.

Assumption A2 is analogous to the assumption of simultaneous treatment and homogeneous treatment effects in a canonical difference-in-differences model (Card and Krueger, 1993; Angrist and Pischke, 2009). Applied to our setting, A2 implies that #MeToo treats all adjudicators once and suddenly and that treatment effects for always reporters do not grow over time. Using the notation of our theoretical model, this assumption states that there is a constant evidentiary threshold $\underline{e}^0$ after October 2017. A2 is plausible if adjudicators do not observe case filing time, or are disallowed from considering it when making their determination. To add credibility to this assumption, we plot the win rate for a subset of always reporters in the first six months after #MeToo and, in line

with A2, show that the win rate remains stable in this window (see Appendix B.4).

The following proposition states our identification result.

Proposition 2. *Under assumptions A1 and A2,* $\beta^O = ATT^{AR}$.

As previously described, A1 allows β^O to identify the ATT for complainants in the overlap group in a similar way that the standard parallel trends assumption allows β^{CE} to identify ATT^{CE}. A2 allows for extrapolation to the entire set of always reporters, i.e., including those who file after #MeToo and therefore do not directly contribute to β^O.[23] A formal proof of Proposition 2 is straightforward, and therefore omitted. As described in Section 4, we can similarly identify gender-specific treatment effects, ATT_M^{AR} and ATT_F^{AR}, under analogues of assumptions A1 and A2 that apply specifically to men and women, as the coefficients of interest from our main gender regressions in Equation 2a and 2b applied to the overlap sample. We refer to these coefficients as β_M^O and β_F^O, respectively.

One point of possibly independent interest is the economic interpretation of ATT^{IR}. While its definition, $\mathbb{E}[Y_{i0}(1) - Y_{i-1}(0)|D_{i0} = 1, IRs] - \mathbb{E}[Y_{i0}(0) - Y_{i-1}(0)|D_{i0} = 1, IRs]$, is mathematically coherent, (potential) outcomes for induced reporters are not observed in the pre-period. However, the standard parallel trends assumption and A1 combine to give the following characterization of ATT^{IR}, and its proof is contained in Appendix A.

Remark 2. *Under the standard parallel trends assumption and A1, the untreated potential outcomes of IRs equals the untreated potential outcomes of the control group, yielding* $ATT^{IR} = \mathbb{E}[Y_{i0}(1) - Y_{i-1}(0)|D_{i0} = 1, IRs] - \mathbb{E}[Y_{i0}(0) - Y_{i-1}(0)|D_{i0} = 0]$.

First, this result allows us in Section 8 to interpret ATT^{IR} similarly to the other treatment effects we study: how would the case outcomes of induced reporters have changed compared to that of control-group reporters who were unaffected by #MeToo? Second, this result allows for direct estimation of ATT^{IR} in an alternative setting and empirical strategy where induced reporters, as opposed to always reporters, can be isolated.

Having established β^O identifies ATT^{AR}, we turn to estimation. The estimates are presented in

[23] Note that Proposition 2 considers a classic two-period difference-in-differences setting. However, this is merely for notational convenience, as the two-period ATT is numerically equivalent to a multi-period ATT given the absence of staggered treatment timing in our setting.

Table 3, and the dynamic analogs are shown in Figures 5a and 5b. In Table 3, the first two columns present $\hat{\beta}^O$, and the third column presents $\hat{\beta}^O_M$ (top row) and $\hat{\beta}^{DDD} = \hat{\beta}^O_W - \hat{\beta}^O_M$ (bottom row). The estimate for overlap cases is $\hat{\beta}^O = 0.087$ (std. err. 0.013), which is positive and significant at the 1% level. There is gender heterogeneity: $\hat{\beta}^O_M = 0.153$ is significantly larger than $\hat{\beta}^O_W = 0.086$. Figure 5a presents a visual lack of pretrends and a positive effect post-#MeToo; 5b presents a visual lack of pretrends for both genders, and a larger increase in the win rate for men than for women.[24] We show heterogeneity by state in Appendix B.3, and results for secondary outcomes, such as the rates of settlement and exiting to court, in Appendix B.5.

We next conduct several robustness checks that validate our use of overlap cases in this exercise, providing empirical evidence showing overlap cases do not systematically differ from the broader set of always reporter cases.[25]

The first potential threat to overlap representativeness is selection on resolution timing. If long complaints are mechanically more likely to resolve after #MeToo (thereby being classified as overlap) than short complaints, and if duration is correlated with outcome, is our overlap sample biased by over-representation of long cases? Our analysis suggests that this concern is likely to be small. If long versus short (and founded versus unfounded) cases are filed at a given rate in the pre-period, then given a sufficiently long time window before #MeToo, the distribution of time since filing for pending cases in each pre-period should be roughly the same.[26] To support this claim, we plot the average duration of open cases in the months leading up to #MeToo in Figure A8. The average case duration is relatively consistent, for both sexual harassment and non-sexual harassment cases, suggesting that overlap cases are unlikely to be selected on duration.

The second threat to overlap representativeness is that case filings may be seasonal—driven by quarterly firm timelines or filing slow down before the year-end—making overlap complaints systematically different from pre-#MeToo cases. Such seasonality would likely be difficult to observe in our event-study plots, which aggregate coefficients to the yearly level. We find that

[24] In both panels, estimated effects are larger in $t = 1$ than in $t = 0$. Notice that the probability of case resolution time in any period $t \geqslant 0$ for overlap complaints need not be independent of case characteristics (either foundedness or time of filing). The observed dynamic pattern post-#MeToo is plausibly caused by non-random resolution timing of cases, and thus, there is no observable violation of assumption A2.

[25] Thematically, these robustness checks test whether the link between Proposition 1 and Remark 1 holds empirically, i.e., are the overlap complainants indeed *unselected* using our formal definition?

[26] Note that this claim applies only to the pre-period, i.e., we do not require #MeToo to have no effect on case duration.

overlap cases *do* exhibit seasonality: they are mechanically less likely to be filed in winter (due to our sample restriction on resolution timing) and most likely to be filed in summer. We test the sensitivity of our results to filing season and show that the sign and relative magnitudes remain broadly consistent across seasons. Even when restricting to winter-filed overlap cases only, point estimates stay in the same ballpark as the pooled results in Table 3. In other words, seasonality of overlap complaint filing is unlikely to explain the patterns in Table 3. For more discussion of robustness to seasonality, see Appendix B.4.

Finally, the third threat to overlap representativeness is that 2017—which all overlap cases necessarily "straddle"— importantly differs from other years for reasons other than #MeToo. To address this, we compare actual 2017 overlap cases to placebo 2016 overlap cases (i.e., cases that would have overlapped had #MeToo occurred in 2016). We find few meaningful differences in case characteristics such as filing season or gender composition between the real and placebo treatment year, adding credibility to the claim that actual overlap cases are not systematically different. For more discussion, see Appendix B.6.

7.2 Estimating the share of always reporters, ω

Having estimated ATT^{CE} and ATT^{AR}, in this section we briefly describe the role of ω.

ω is a latent variable that depends on the number of cases reported before and after #MeToo, and the unobserved distribution of foundedness for reporters both before and after #MeToo (which itself is potentially affected by the rate of realized incidence of sexual harassment). As the second component is difficult to observe directly, estimating ω precisely is difficult in our context.

To provide some guidance to the reader, we calibrate two plausible values—$\omega_1 = 0.801$ and $\omega_2 = 0.928$—using assumptions and estimates of sexual harassment incidence from prior literature (Levy and Mattsson, 2025). The derivation of these calibrated values is detailed in Appendix D.

However, our qualitative takeaways on the directions of the treatment and selection effect are not sensitive to the value of ω. As we discuss in the upcoming section, our directional takeaways depend on the sign of ATT^{AR} (direct treatment effect) and on an ordinal comparison of ATT^{AR} and ATT^{IR} (selection effect). First, notice that ATT^{AR} mechanically does not depend on ω, meaning that the direct treatment effect is not affected by the share of always reporters. Second, the following

ordinal comparisons of ATT^{AR} and ATT^{IR}—which do not depend on ω—hold by inspection of Equation 4, meaning that the sign of the selection effect will not depend on ω:[27]

Remark 3. *Fix ATT^{CE} and ATT^{AR}. Then regardless of ω:*

- $ATT^{AR} > ATT^{IR}$ *if* $ATT^{AR} > ATT^{CE}$,
- $ATT^{AR} = ATT^{IR}$ *if* $ATT^{AR} = ATT^{CE}$, *and*
- $ATT^{AR} < ATT^{IR}$ *if* $ATT^{AR} < ATT^{CE}$.

8 Selection versus Treatment Decomposition Results

We now synthesize the various theoretical and empirical results presented above to answer our two questions of interest: do adjudicators believe complainants more as a result of #MeToo (direct treatment effect), and does #MeToo increase the share of founded accusations (selection effect)? To ensure identification of the relevant treatment effects (see Sections 4 and 7), we suppose that the *standard parallel trends assumption*, A1, and A2 hold throughout this section.

8.1 Direct Treatment Effect

In Proposition 1, we proved that the direct treatment effect is revealed by the change in the win rate for always reporters after #MeToo, or ATT^{AR}. That is, Proposition 1 states that adjudicators lower their evidentiary threshold if and only if the always reporter win rate increases. As shown in the previous section, ATT^{AR} is identified by β^O, and represents the change in the win rate of sexual harassment always reporters compared to that of always reporters for control-group complaints under our maintained assumptions.

We find evidence that adjudicators lower their evidentiary thresholds in response to #MeToo, i.e., that they "believe" more accusations. This follows because our estimate of ATT^{AR}, $\hat{\beta}^O$, is positive and significant ($\hat{\beta}^O = 0.087$ and is statistically different from zero at the 1% level, see col. 1 of Table 3). In Figure 6a we plot ATT^{AR} across values of ω.

We find adjudicators "believe" more complainants of both genders. For women $\hat{\beta}^O_W = 0.153 -$

[27] As in Remark 1, the findings of the current remark are unchanged if we replace ATT^{AR} with the ATT for any unselected subset of complainants, and ATT^{IR} with the ATT of the complement of the unselected subset.

$0.067 = 0.086$, which is again statistically different from zero at the 1% level (see col. 3 of Table 3). For men, the increase is even greater. $\hat{\beta}_M^O = 0.153$ which is both statistically significant at the 1% level and larger than the analogous effect for women. Figure 6b plots these effects by gender.

Interpreting these results in light of Proposition 1 yields the following conclusions:

Finding 1. *#MeToo induced adjudicators to lower the evidentiary threshold for believing sexual harassment complainants for both genders. The effect is larger, by roughly 50%, for men than for women, i.e., adjudicators "believe" sexual harassment complainants of both sexes more after #MeToo, but especially those of male complainants.*

8.2 Selection Effect

Proposition 1 finds that the selection effect is revealed by comparing the win rates of always versus induced reports after #MeToo. Induced reporters are positively selected (for foundedness) if and only if the win rate is higher for induced reporters than always reporters. However, two complications arise when attempting to measure these quantities directly. First, we do not observe which subsample of complaints forms the always reporter set, or even what the relative size ω of the always reporters is—we know only that overlap complainants are unselected. Second, due to the endogeneity of reporting decisions and the mechanical timing bias of overlap cases, directly comparing win rates may misrepresent the direction or magnitude of the effect.[28]

To sidestep these timing issues, we proceed by studying the selection effect by comparing treatment effects for always and induced reporters. Of course, the sign of the difference in *treatment effects*—i.e., $ATT^{IR} - ATT^{AR}$—may not be the same as the sign of the difference in *win rates*—i.e.,

[28]There are a number of complications when using the win rates in the post-#MeToo period to directly answer this question. First, our unselected sample of complaints, the overlap complaints, happen immediately post-#MeToo by construction. Second, our theoretical model does not speak to the following reality: positively versus negatively selected cases may not be filed uniformly at random after October 2017 because the timing of filing a report is endogenously chosen by the complainant. If filing time and case foundedness are correlated, we may naively infer that induced reporters were positively selected for foundedness if we looked at the immediate post-#MeToo period, but that they were negatively selected if we looked farther into the future (or vice versa). For example, consider two potential reporters who decide they want to file complaints immediately after #MeToo. Reporter 1 experienced a founded occurrence of sexual harassment in the pre-period, but chose not to make a report at the time. However, given the increase in (perceived) win probability post #MeToo, she is incentivized to do so immediately. By contrast, Reporter 2 never experienced sexual harassment, but decides he wants to file an unfounded report now that he perceives a higher chance of winning, and files his accusation several months or years after #MeToo. The generality of our model—recall that the selection effect also nests the effect of deterrence (and other factors that change the foundedness distribution of cases), which may take time to affect the actual incidence of sexual harassment—leads to even more complications when trying to directly compare win rates across time.

$\mathbb{E}[Y_{i0}(1)|D_{i0} = 1, IR] - \mathbb{E}[Y_{i0}(1)|D_{i0} = 1, AR]$—generally. To address this possibility, the following proposition states an equivalence between the ATT and post-period win rates of always versus induced reporters: under an additional assumption, ATT^{IR} is higher than ATT^{AR} if and only if post-period win rate of induced reporters is higher than the post-period win rate of always reporters.

Proposition 3. *Suppose standard parallel trends, A1, and the following* No Extreme Baseline Gap in Potential Outcomes *condition hold:* $\left|\mathbb{E}[Y_{i-1}(0)|D_{i0} = 1, IR] - \mathbb{E}[Y_{i-1}(0)|D_{i0} = 1, AR]\right| \leqslant \left|\mathbb{E}[Y_{i0}(1)|D_{i0} = 1, IR] - \mathbb{E}[Y_{i0}(1)|D_{i0} = 1, AR]\right|$*, or that the gap in (untreated) potential outcomes between IRs and ARs is magnified by #MeToo. Then* $ATT^{IR} \geqslant ATT^{AR}$ *if and only if* $\mathbb{E}[Y_{i0}(1)|D_{i0} = 1, IR] \geqslant \mathbb{E}[Y_{i0}(1)|D_{i0} = 1, AR]$*, or that the post-#MeToo win rate of IR sexual harassment complaints exceeds that of the post-#MeToo win rate of AR sexual harassment complaints.*

The no extreme baseline gap condition is satisfied when the evidentiary bar for winning is sufficiently high pre-#MeToo and is lowered as a result of #MeToo.[29] Naturally, the conclusions of Proposition 3 hold within gender as well, that is, conditioning the no extreme baseline gap restriction on a gender g implies that $ATT_g^{IR} \geqslant ATT_g^{AR}$ if and only if the post-#MeToo win rate of IR sexual harassment complaints by gender g exceeds that of the post-#MeToo win rate of AR sexual harassment complaints by gender g. Recall that the benefit of considering treatment effects rather than raw win rates is that our estimates are valid regardless of the endogeneity of filing timing after #MeToo. However, in Appendix B.7 we conduct regression discontinuity analyses on raw win rates immediately in the wake of #MeToo that corroborate our gender-specific findings below.

To address the issue of ω being a latent variable, we consider the point estimates of treatment effects $\hat{\beta}^{O}$ (which identifies ATT^{AR}) and $\hat{\beta}^{CE}$ (which identifies the combined ATT) as a function of ω to reveal the ATT of induced reporters by using Equation 4. Doing so allows for a clean characterization of the relative sizes of the selection and direct treatment effects for different values of ω. Intuitively, when ω is small the direct treatment effect dominates, and when ω is large the selection effect dominates.

[29] This claim follows from our theoretical framework presented in Section 6. The smoothness conditions on distributions G_F and G_U imply that for a sufficiently high evidentiary threshold in the pre-period ($\underline{e}^{-1} \approx 1$), the win rate for both founded and unfounded cases is $1 - G_F(\underline{e}^{-1}) \approx 1 - G_U(\underline{e}^{-1}) \approx 0$, implying that the baseline gap in potential outcomes in the pre-period converges to 0 as the evidentiary threshold converges to 1.

We find evidence that the selection effect of #MeToo is positive, that is, induced reporters bring forth founded claims of sexual harassment. However, there is significant gender heterogeneity.

Pooling across genders, we see that $\hat{\beta}^{CE} = 0.101$ exceeds $\hat{\beta}^{O} = 0.087$ (see col. 1 of Tables A6 3), i.e., that ATT^{IR} exceeds ATT^{AR} (see Remark 1 and Proposition 2). In Figure 6a, we trace out the treatment effects for induced reporters for different values of ω using Equation 4. The values of ω we calibrated in Section 7.2 are labeled ω_1 and ω_2. The treatment effects for induced reporters exceed those for always reporters under every value of ω (as predicted by Remark 3). For women, the pattern is the same as the aggregate effects: $\hat{\beta}^{CE}_{W} = 0.119 + 0.004 = 0.123$ is larger than $\hat{\beta}^{O}_{W} = 0.153 - 0.067 = 0.086$ (see col. 3 of Tables 2 and 3). Figure 6b shows that treatment effect for female induced reporters exceeds the treatment effect for female always reporters for all values of ω (orange lines). For men, $\hat{\beta}^{CE}_{M} = 0.119$ is smaller than $\hat{\beta}^{O}_{M} = 0.153$ (see col. 3 of Tables 2 and 3). The treatment effect for male induced reporters is *lower* than that for male always reporters (blue lines of Figure 6b) for all values of ω. Combining these observations with Propositions 1 and 3 yields the following conclusions:

Finding 2. *#MeToo induced more founded reports to be filed. However, there is significant gender heterogeneity in the foundedness of induced reporters: women are positively selected for foundedness, while men are negatively selected for foundedness.*

Indeed, for large values of ω, the treatment effect for male induced reporters is actually *negative*. This means that for our calibrated values of ω, the negative selection effect for men is so large that it dominates the positive adjudicator treatment effect. Moreover, for our calibrated values of ω, the selection effect is larger than the direct treatment effect for women, although both of these effects push in the same direction (i.e., increase women's win rates). Viewing these findings in light of Remark 2 implies:

Finding 3. *If the share of always reporters ω exceeds 0.778—our calibrated values are $\omega_1 = 0.801$ and $\omega_2 = 0.928$—male induced reporters have a lower win rate than male always reporters in a counterfactual world where #MeToo never happened. In other words, men induced to report by #MeToo won less often than men who reported before #MeToo, despite benefiting from procedural advantage. Moreover, if ω exceeds 0.570, women induced to report by #MeToo had a treatment effect twice as large as women always reporters.*

Note, as previously discussed, we investigate the robustness of our results to alternative samples, alternative specifications, and alternative control groups in Appendix C. Broadly speaking, our main empirical takeaways are qualitatively and quantitatively similar across these alternatives.

9 Conclusion

In this paper, we study the #MeToo movement, which aimed to increase the reporting of sexual harassment, and evaluate a common criticism that the movement encouraged the reporting of unfounded accusations. First, we quantify the total impact of the #MeToo movement on sexual harassment complaints. Our baseline difference-in-differences estimates find that #MeToo benefited complainants by increasing the probability that they win their case, and that it did so equally for male and female complainants. However, this effect nests both changes in how much adjudicators "believe" complainants and changes in the composition of complaints reported.

Second, to disentangle treatment and selection effects, we present a novel decomposition of the difference-in-differences estimate into an effect on always reporters (reporters whose composition of case fundamentals match those of reporters before #MeToo) and an effect on induced reporters (those encouraged to report by #MeToo who may be positively or negatively selected). Intuitively, these two groups reveal, respectively, the treatment effect and the combination of the treatment and selection effects. We estimate the effect on always reporters using "overlap" complaints filed before #MeToo but resolved after. We find that the probability of winning for induced reporters is higher than always reporters for women, but that the reverse is true for men.

We interpret our empirical results through a simple model, which suggests that #MeToo had a dual effect. First, adjudicators were more likely to "believe" complaints after #MeToo, holding the characteristics of the complaint fixed, and this effect was larger for men than for women. More formally, we find evidence that adjudicators lowered their evidentiary threshold for findings of probable cause that sexual harassment occurred. Second, women induced to report by #MeToo were more likely to file a founded complaint than women who would have reported in the absence of #MeToo. On the other hand, men induced to report by #MeToo were less likely to file a founded complaint than men who would have reported in the absence of #MeToo.

These findings contribute to our broader understanding of social campaigns and movements that increase the reporting of wrongdoing, highlighting their dual effect on how adjudicators weigh evidence and on the composition of people who choose to report as a result.

References

Abi Adams-Prassl, Kristiina Huttunen, Emily Nix, and Ning Zhang. Violence against women at work. *The Quarterly Journal of Economics*, 139(2):937–991, 2024.

Monice Anderson and Skye Toor. How social media users have discussed sexual harassment since #MeToo went viral. https://www.pewresearch.org/short-reads/2018/10/11/how-social-media-users-have-discussed-sexual-harassment-since-metoo-went-viral/, 2018. Accessed: 2024-08-09.

Joshua D Angrist and Jörn-Steffen Pischke. *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press, 2009.

David Arnold, Will Dobbie, and Crystal S Yang. Racial bias in bail decisions. *The Quarterly Journal of Economics*, 133(4):1885–1932, 2018.

Elliott Ash, Daniel L Chen, and Suresh Naidu. Ideas have consequences: The impact of law and economics on american justice. *The Quarterly Journal of Economics*, 141(1):845–887, 2026.

Cyprien Batut, Caroline Coly, and Sarah Schneider-Strawczynski. *It's a man's world: culture of abuse, #MeToo and worker flows*. Università Bocconi, The Dondena Centre, 2021.

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

Dan A Black, Joonhwi Joo, Robert LaLonde, Jeffrey A Smith, and Evan J Taylor. Simple tests for selection: Learning more from instrumental variables. *Labour Economics*, 79:102237, 2022.

Laura Boudreau, Sylvain Chassang, Ada González-Torres, and Rachel Heath. Monitoring harassment in organizations. 2024.

Xiqian Cai, Shuai Chen, Zhengquan Cheng, and Emily E. Nix. Gender-based violence and judge responses. 2025.

Sebastian Calonico, Matias D Cattaneo, and Rocio Titiunik. Robust data-driven inference in the regression-discontinuity design. *The Stata Journal*, 14(4):909–946, 2014.

Travis Campbell. Black lives matter's effect on police lethal use of force. *Journal of Urban Economics*, 141:103587, 2024.

David Card and Alan B Krueger. Minimum wages and employment: A case study of the fast food industry in new jersey and pennsylvania. *The American Economic Review*, 84(4):772–793, 1993.

Sylvain Chassang and Gerard Padró i Miquel. Crime, Intimidation, and Whistleblowing: A Theory of Inference from Unverifiable Reports. *The Review of Economic Studies*, 86(6):2530–2553, 2018.

Ing-Haw Cheng and Alice Hsiaw. Reporting sexual misconduct in the #MeToo era. *American Economic Journal: Microeconomics*, 14(4):761–803, 2022.

Gordon B Dahl and Matthew M Knepper. Why is workplace sexual harassment underreported? the value of outside options amid the threat of retaliation. *American Economic Review*, 116(3): 897–933, 2026.

Shai Danziger, Jonathan Levav, and Liora Avnaim-Pesso. Extraneous factors in judicial decisions. *Proceedings of the National Academy of Sciences*, 108(17):6889–6892, 2011.

Tanaya Devi and Roland G Fryer Jr. Policing the police: The impact of "pattern-or-practice" investigations on crime. 2020.

Ozkan Eren and Naci Mocan. Emotional judges and unlucky juveniles. *American Economic Journal: Applied Economics*, 10(3):171–205, 2018.

Olle Folke and Johanna Rickne. Sexual harassment and gender inequality in the labor market. *The Quarterly Journal of Economics*, 137(4):2163–2212, 2022.

Eleanor Frisch. State sexual harassment definitions and disaggregation of sex discrimination claims. *Minn. L. Rev.*, 98:1943, 2013.

Germain Gauthier. Measuring crime reporting and incidence: Method and application to #MeToo. 2025.

Melanie Heenan and Suellen Murray. Study of reported rapes in victoria 2000–2003: Summary research report. *Retrieved from the State of Victoria (Australia), Department of Human Services*, 2006.

Joni Hersch. Compensating differentials for sexual harassment. *American Economic Review P&P*, 101(3):630–634, 2011.

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

Anthony Heyes and Soodeh Saberian. Temperature and decisions: evidence from 207,000 court cases. *American Economic Journal: Applied Economics*, 11(2):238–265, 2019.

Louis Kaplow. On the optimal burden of proof. *Journal of Political Economy*, 119(6):1104–1140, 2011.

Lars J Kirkeboen, Edwin Leuven, and Magne Mogstad. Field of study, earnings, and self-selection. *The quarterly journal of economics*, 131(3):1057–1111, 2016.

David S. Lee. Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *The Review of Economic Studies*, 76(3):1071–1102, 2009.

Legal Information Institute. Sexual harassment. https://www.law.cornell.edu/wex/sexual_harassment. Accessed: 2024-08-09.

Ro'ee Levy and Martin Mattsson. The effects of social movements: Evidence from #MeToo. Available at SSRN 3496903, 2025.

David Lisak, Lori Gardinier, Sarah C Nicksa, and Ashley M Cote. False allegations of sexual assault: An analysis of ten years of reported cases. *Violence against women*, 16(12):1318–1334, 2010.

Kimberly A Lonsway and Joanne Archambault. The "justice gap" for sexual assault cases: Future directions for research and reform. *Violence against women*, 18(2):145–168, 2012.

Jack Mountjoy. Community colleges and upward mobility. *American Economic Review*, 112(8): 2580–2630, 2022.

National Women's Law Center. State bills related to #MeToo. https://nwlc.org/wp-content/uploads/2022/10/MeToo-State-Bills-2023-10.18.23.pdf, 2023. Accessed: 2024-08-09.

Andreas Olden and Jarle Møen. The triple difference estimator. *The Econometrics Journal*, 25(3): 531–553, 2022.

Pew Research Center. Sexual harassment at work in the era of #MeToo. https://www.pewresearch.org/social-trends/2018/04/04/sexual-harassment-at-work-in-the-era-of-metoo/, 2018. Accessed: 2024-08-09.

A. Raj, N. Rao, P. Patel, H. Kearl, and C. Skolnick. #MeToo 2024: A National Study of Sexual Harassment and Assault in the United States. 2024.

Stop Street Harassment. 2019 #MeToo national sexual harassment and assault report. https://stopstreetharassment.org/wp-content/uploads/2012/08/2019-MeToo-National-Sexual-Harassment-and-Assault-Report.pdf, 2019. Accessed: 2024-08-09.

U.S. Bureau of Labor Statistics. Labor force participation rate for women highest in the district of columbia in 2022. https://www.bls.gov/opub/ted/2023/labor-force-participation-rate-for-women-highest-in-the-district-of-columbia-in-2022.htm, 2023. Accessed: 2025-08-04.

U.S. Bureau of Labor Statistics. Highlights of women's earnings in 2022. https://www.bls.gov/opub/reports/womens-earnings/2022/, 2024. Accessed: 2025-08-04.

Tables & Figures

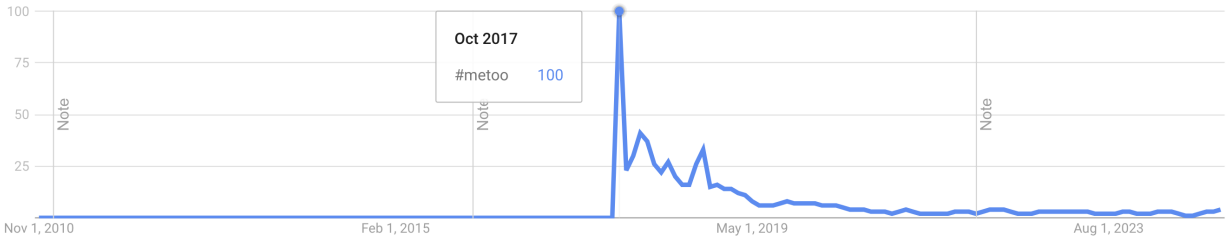


Figure 1: Google Trends timeseries of #MeToo searches from 2010 to 2024, showing a peak in October 2017. The y-axis is Google's measure of the relative popularity of the search term over time, with 100 being peak popularity.

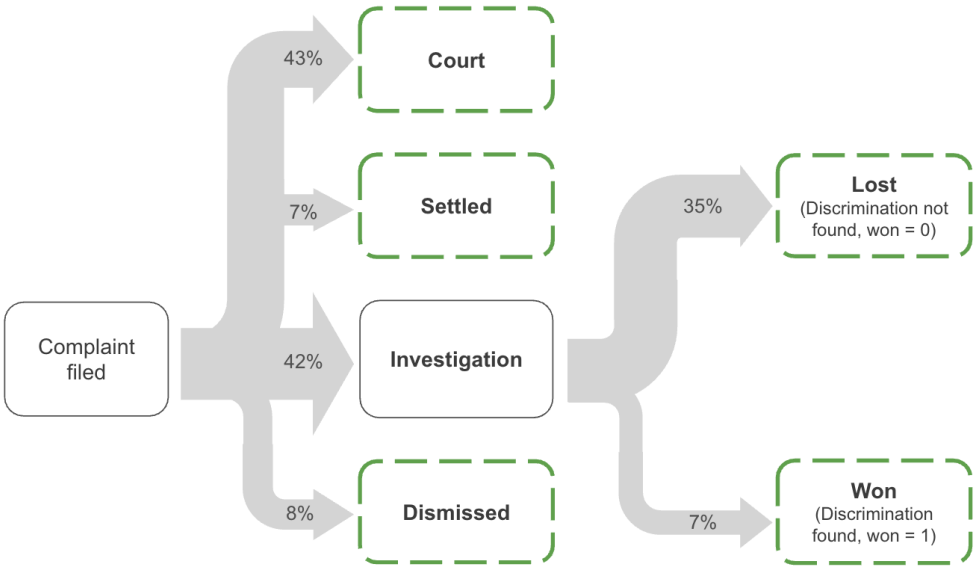


Figure 2: Complaint flow with terminal outcomes in boxes with green dashed lines. All terminal boxes add to 100%. Won/Lost are unconditional shares of all filed complaints. Conditional win rate among investigated cases equals Won/(Won+Lost).

TABLE 1 — COMPLAINT CHARACTERISTICS

	Sample Statistics		Mean Difference (Post-Pre) MeToo	
	All	SH only	All	SH only
Complaint Characteristics				
Sexual harassment	0.171	1.000	-0.096***	.
	[0.38]	[0.00]	(.0019)	.
Complainant is female	0.499	0.741	-0.027***	0.041***
	[0.50]	[0.44]	(.004)	(.0098)
Filed after MeToo	0.724	0.837	.	.
	[0.45]	[0.37]	.	.
Duration (days)	162.583	117.125	65.204***	150.072***
	[176.71]	[167.06]	(.89)	(2.4)
Discrimination basis				
Age	0.074	.	0.049***	.
	[0.26]	.	(.0013)	.
Disability	0.337	.	-0.038***	.
	[0.47]	.	(.0024)	.
Nationality	0.030	.	0.024***	.
	[0.17]	.	(.00088)	.
Race	0.254	.	0.098***	.
	[0.44]	.	(.0022)	.
Religion	0.027	.	-0.008***	.
	[0.16]	.	(.00082)	.
Retaliation	0.110	0.006	-0.026***	0.019***
	[0.31]	[0.08]	(.0016)	(.0012)
Sex	0.168	0.986	-0.098***	-0.030***
	[0.37]	[0.12]	(.0019)	(.0018)
Outcomes				
Settled	0.068	0.059	0.043***	0.082***
	[0.25]	[0.24]	(.0013)	(.0035)
Won	0.066	0.082	0.092***	0.216***
	[0.25]	[0.27]	(.0013)	(.0039)
Lost	0.355	0.184	0.240***	0.268***
	[0.48]	[0.39]	(.0024)	(.0056)
Dismissed	0.074	0.063	0.047***	0.094***
	[0.26]	[0.24]	(.0014)	(.0036)
Went to court	0.438	0.609	-0.418***	-0.657***
	[0.50]	[0.49]	(.0024)	(.0063)
Observations	191,985	32,737	191,985	32,737

Table 1: SH is sexual harassment. Columns 1 and 2 present the means and standard deviations in parentheses of variables in the entire sample and the sample of sexual harassment complaints. Won/Lost are unconditional shares of all filed complaints. Conditional win rate among investigated cases equals Won/(Won+Lost). Columns 3 and 4 present the difference in means of each variable after minus before #MeToo, with the standard error of the difference-in-means test in parentheses. Sexual harassment can be classified as Retaliation in the data, so Retaliation is non-zero in the "SH only" column.

FIGURE 3a — IMPACT OF #METOO ON PROBABILITY COMPLAINANT WINS

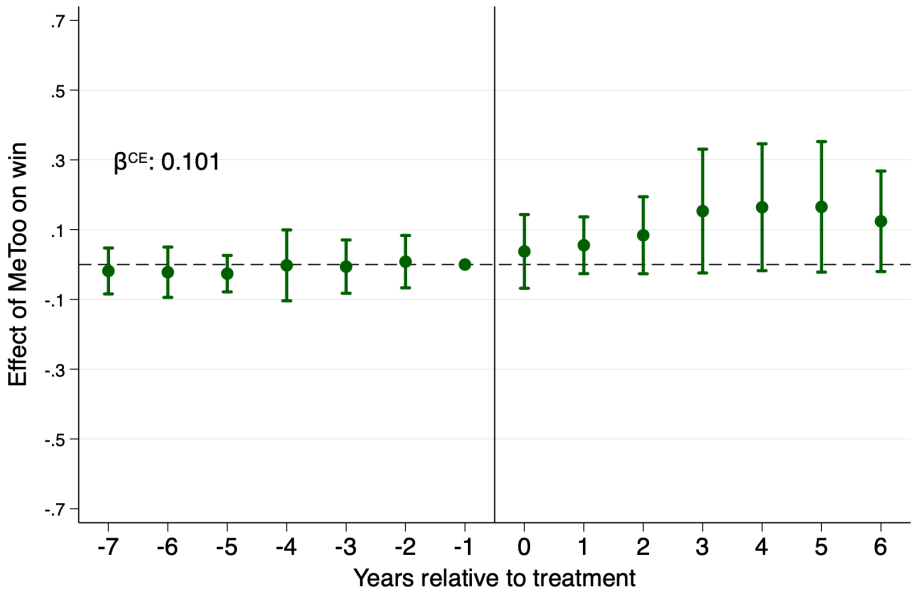


FIGURE 3b — IMPACT OF #METOO ON PROBABILITY COMPLAINANT WINS, BY GENDER

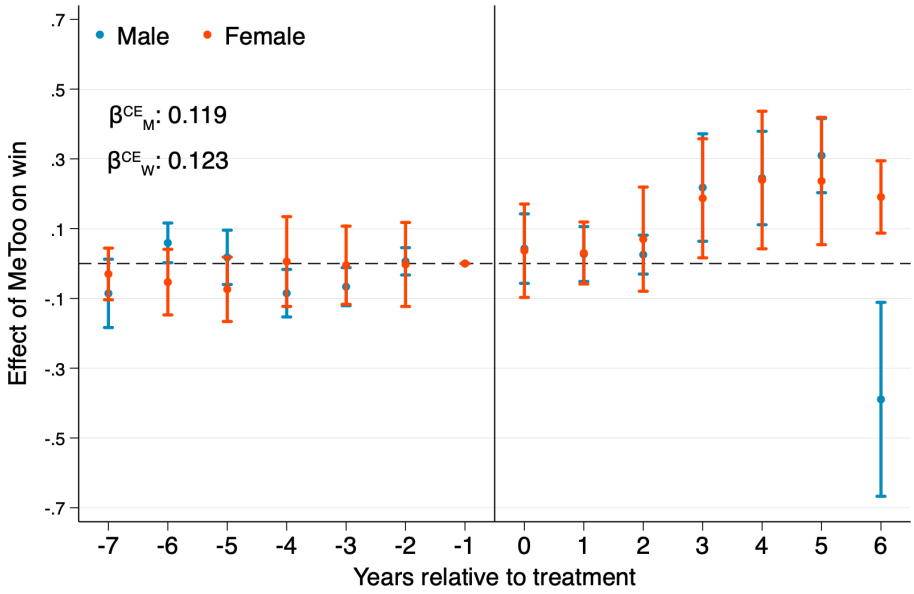


Figure 3: These figures provide estimates of Equations 1b and 2b. The x-axis in all subfigures represents 12-month intervals centered around the #MeToo start date. $t = -1$ is the 12 months preceding October 15, 2017, $t = 0$ is the 12-month period following October 15, 2017, and so on. In the top panel, we control for complaint type-by-state and year-month-by-state fixed effects, where the "year-month" specifies the date of the complaint resolution. In the bottom panel, we additionally interact these same fixed effects with complainant gender. 95% confidence intervals are reported using standard errors clustered at the complaint type by state level.

TABLE 2 — EFFECT OF #METOO ON LIKELIHOOD COMPLAINANT WINS SEXUAL HARASSMENT CASE

	All complaints	**Complaints with gender**	
SH × Post	0.101***	0.116***	0.119***
	(0.029)	(0.027)	(0.026)
SH × Post × Female			0.004
			(0.022)
Unit and Time × State FE	✓	✓	✓
Unit and Time × State × Female FE			✓
N	73,995	46,750	46,746
R^2	0.171	0.089	0.096
Control mean	0.152	0.226	0.229

Table 2: SH is sexual harassment. Columns 1 and 2 present estimates of Equation 1a. Column 3 presents estimates of Equation 2a. Fixed effects are complaint type by state and year-month of case resolution by state. For the triple differences specification, the fixed effects are complaint type by state by female complainant and year-month of case resolution by state by female complainant. Standard errors clustered at the state by complaint type level. Single asterisk denotes statistical significance at the 90% level of confidence; double, 95%; triple, 99%.

FIGURE 4 — ILLUSTRATION OF 2×2 COMPARISON IN OVERLAP REGRESSIONS

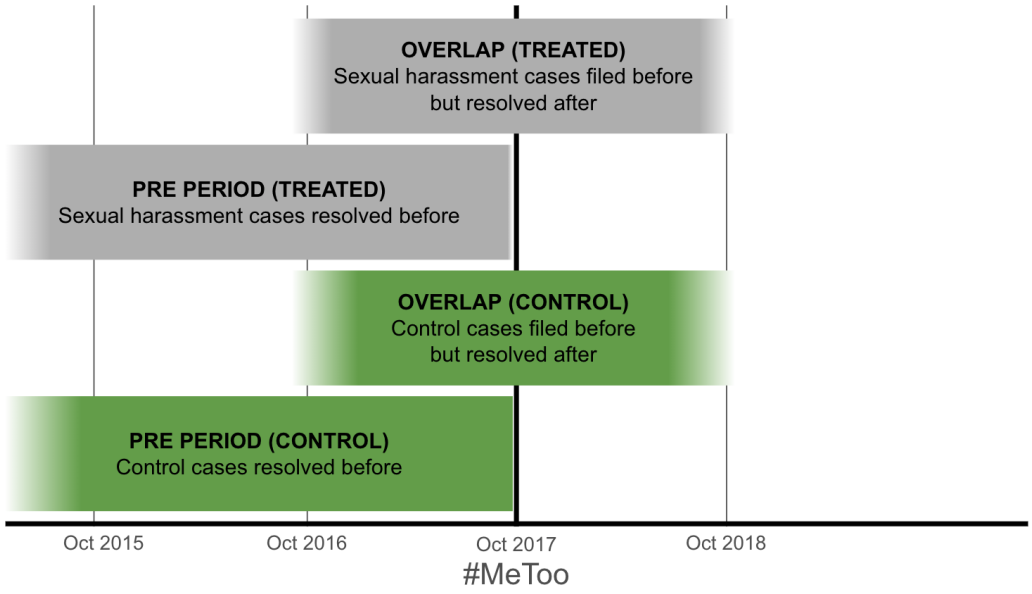


Figure 4: Stylized illustration of the two-by-two comparison that our difference-in-differences regression for overlap cases is making.

FIGURE 5a — IMPACT ON WINNING, OVERLAP

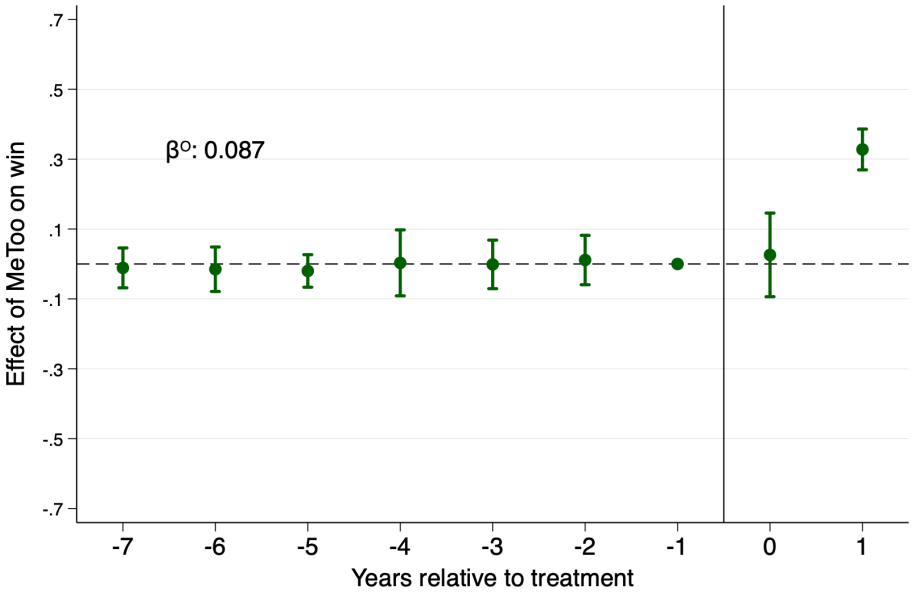


FIGURE 5b — IMPACT ON WINNING BY GENDER, OVERLAP

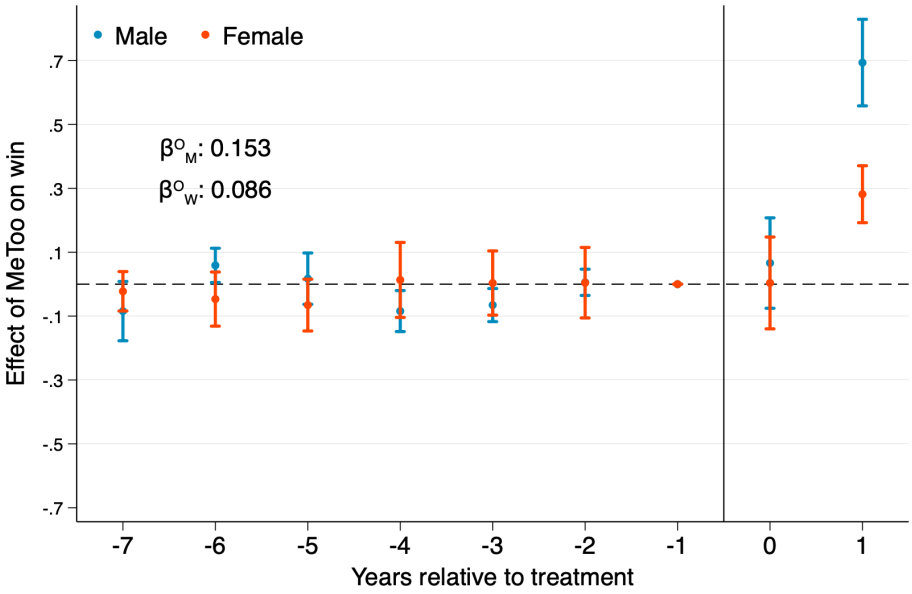


Figure 5: These figures provide estimates of Equations 1b and 2b for the subset of cases that are filed before #MeToo (the overlap sample from Figure 4). The x-axis in all subfigures represents 12-month intervals centered around the #MeToo start date. $t = -1$ is the 12 months preceding October 15, 2017, $t = 0$ is the 12-month period following October 15, 2017, and so on. We control for unit-by-state and year-month-by-state fixed effects, where the "year-month" specifies the date of the complaint resolution. 95% confidence intervals are reported using standard errors clustered at the complaint type by state level.

FIGURE 6a — ATT^{AR} AND ATT^{IR} BY SHARE OF ALWAYS REPORTERS

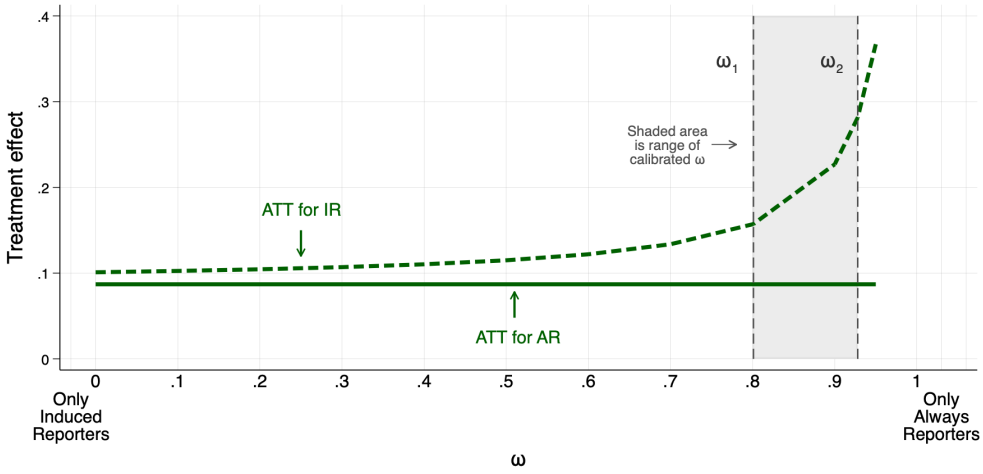


FIGURE 6b — ATT^{AR} AND ATT^{IR} BY GENDER

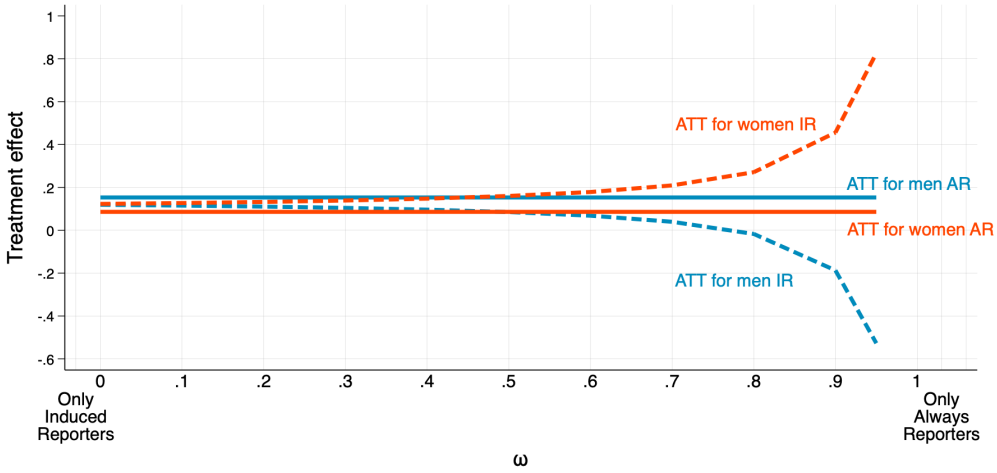


Figure 6: In the top panel, we indicate ω_1 and ω_2 from Section 7.2 with vertical lines. IR refers to induced reporters and AR refers to always reporters. The treatment effect for AR is constant across all values of ω. The treatment effect for IR asymptotes at $\omega = 1$, so the graphs are bounded at $\omega = 0.95$.

TABLE 3 — EFFECT OF #METOO ON LIKELIHOOD COMPLAINANT WINS SEXUAL HARASSMENT CASE IN OVERLAP SAMPLE

	All complaints	**Complaints with gender**	
SH × Post	0.087***	0.102***	0.153***
	(0.013)	(0.011)	(0.014)
SH × Post × Female			-0.067**
			(0.027)
Unit and Time × State FE	✓	✓	✓
Unit and Time × State × Female FE			✓
N	39,321	30,741	30,733
R^2	0.128	0.073	0.081
Control mean	0.152	0.226	0.229

Table 3: SH is sexual harassment. Columns 1 and 2 present estimates of Equation 1a estimated only on complaints filed before #MeToo (see Figure 4). Column 3 presents estimates of Equation 2a estimated only on complaints filed before #MeToo. Fixed effects are basis of discrimination by state and year-month of case resolution by state. For the triple differences specification, the fixed effects are basis of discrimination by state by female complainant and year-month of case resolution by state by female complainant. Standard errors clustered at the state by complaint type (unit) level. Single asterisk denotes statistical significance at the 90% level of confidence; double, 95%; triple, 99%.

Appendix

A Proofs

Proof of Proposition 1

Proof of Part 1. Let r^0 be a set of always reporters. Recall that the win rate is defined only for cases that go through an investigation, and that court and settlement outcomes are independent of a case's foundedness. Therefore, the win rate among always reporters in each period t is given by:

$$W^t := \frac{|r_F^0| \cdot Pr(e \geqslant \underline{e}^t | \tau = F) + |r_U^0| \cdot Pr(e \geqslant \underline{e}^t | \tau = U)}{|r_F^0| + |r_U^0|} = \frac{|R_F^{-1}| \cdot Pr(e \geqslant \underline{e}^t | \tau = F) + |R_U^{-1}| \cdot Pr(e \geqslant \underline{e}^t | \tau = U)}{|R_F^{-1}| + |R_U^{-1}|}$$

where the equality follows from the definition of always reporters. By the properties of probability distributions, $Pr(e \geqslant \underline{e}^0 | \tau = F) \geqslant Pr(e \geqslant \underline{e}^{-1} | \tau = F)$ and $Pr(e \geqslant \underline{e}^0 | \tau = U) \geqslant Pr(e \geqslant \underline{e}^{-1} | \tau = U)$ if and only if $\underline{e}^0 \leqslant \underline{e}^{-1}$. Therefore, $W^0 \geqslant W^{-1}$ if and only if $\underline{e}^0 \leqslant \underline{e}^{-1}$, as desired. □

Proof of Part 2. This follows straightforwardly using the machinery from the proof of part 1. Because the same evidentiary threshold $\underline{e}^0$ is used for all cases (i.e., all reporter types) post-#MeToo, and because we assume that G_F dominates G_U in the likelihood ratio order, it follows that $Pr(e \geqslant \underline{e}^0 | \tau = F) \geqslant Pr(e \geqslant \underline{e}^0 | \tau = U)$ for any given $\underline{e}^0 \in [0, 1]$. □

Proof of Remark 2

Proof. First, note that the *standard parallel trends assumption* is that

$$\mathbb{E}[Y_{i0}(0) - Y_{i-1}(0) | D_{i0} = 0] = \mathbb{E}[Y_{i0}(0) - Y_{i-1}(0) | D_{i0} = 1]. \tag{5}$$

Also, Assumption A1 is that

$$\mathbb{E}[Y_{i0}(0) - Y_{i-1}(0) | D_{i0} = 0] = \mathbb{E}[Y_{i0}(0) - Y_{i-1}(0) | D_{i0} = 1, ARs]. \tag{6}$$

Because each reporter in the post period is either in the always reporter group or in the induced reporter group, and ω is the share of ARs, the Law of Iterated Expectations implies that:

$$\begin{aligned} E[Y_{i0}(0) - Y_{i-1}(0) \,|\, D_{i0} = 1] = \omega \, & E[Y_{i0}(0) - Y_{i-1}(0) \,|\, D_{i0} = 1, ARs] \\ & + (1 - \omega) \, E[Y_{i0}(0) - Y_{i-1}(0) \,|\, D_{i0} = 1, IRs] \,. \end{aligned} \tag{7}$$

Then it holds that:

$$\begin{aligned}\mathbb{E}[Y_{i0}(0) - Y_{i-1}(0)|D_{i0} = 0] &= \mathbb{E}[Y_{i0}(0) - Y_{i-1}(0)|D_{i0} = 1] \\ &= \omega\mathbb{E}[Y_{i0}(0) - Y_{i-1}(0)|D_{i0} = 1, ARs] + (1-\omega)\mathbb{E}[Y_{i0}(0) - Y_{i-1}(0)|D_{i0} = 1, IRs] \\ &= \omega\mathbb{E}[Y_{i0}(0) - Y_{i-1}(0)|D_{i0} = 0] + (1-\omega)\mathbb{E}[Y_{i0}(0) - Y_{i-1}(0)|D_{i0} = 1, IRs] \\ &= \mathbb{E}[Y_{i0}(0) - Y_{i-1}(0)|D_{i0} = 1, IRs], \end{aligned} \tag{8}$$

where the first equality follows from Equation 5, the second equality follows from Equation 7, the third equality follows from Equation 6, and the final equality comes from solving $\mathbb{E}[Y_{i0}(0) - Y_{i-1}(0)|D_{i0} = 0] = \omega\mathbb{E}[Y_{i0}(0) - Y_{i-1}(0)|D_{i0} = 0] + (1-\omega)\mathbb{E}[Y_{i0}(0) - Y_{i-1}(0)|D_{i0} = 1, IRs]$.

Then combining the definition of ATT^{IR} and Equation 8 yields the desired conclusion.

□

Proof of Proposition 3

Proof. The standard parallel trends assumption and A1 implies the following, as derived in Equations 6 and 8:

$$ATT^{IR} - ATT^{AR} = \mathbb{E}[Y_{i0}(1) - Y_{i-1}(0)|D_{i0} = 1, IRs] - \mathbb{E}[Y_{i0}(1) - Y_{i-1}(0)|D_{i0} = 1, ARs]. \tag{9}$$

First, we show that the sign of $ATT^{IR} - ATT^{AR}$ reveals whether IRs or ARs have a higher expected win rate in the post period.

Suppose $ATT^{IR} - ATT^{AR} \geqslant 0$. Rearranging Equation 9:

$$\mathbb{E}[Y_{i0}(1)|D_{i0} = 1, IRs] - \mathbb{E}[Y_{i0}(1)|D_{i0} = 1, ARs] \geqslant \mathbb{E}[Y_{i-1}(0)|D_{i0} = 1, IRs] - \mathbb{E}[Y_{i-1}(0)|D_{i0} = 1, ARs]. \tag{10}$$

But then Equation 10 and the no extreme baseline gaps condition are both satisfied only if $\mathbb{E}[Y_{i0}(1)|D_{i0} = 1, IRs] - \mathbb{E}[Y_{i0}(1)|D_{i0} = 1, ARs] \geqslant 0$, proving the desired claim.

Similarly, suppose $ATT^{IR} - ATT^{AR} \leqslant 0$. Rearranging Equation 9:

$$\mathbb{E}[Y_{i0}(1)|D_{i0} = 1, IRs] - \mathbb{E}[Y_{i0}(1)|D_{i0} = 1, ARs] \leqslant \mathbb{E}[Y_{i-1}(0)|D_{i0} = 1, IRs] - \mathbb{E}[Y_{i-1}(0)|D_{i0} = 1, ARs]. \tag{11}$$

But then Equation 11 and the no extreme baseline gaps condition are both satisfied only if $\mathbb{E}[Y_{i0}(1)|D_{i0} = 1, IRs] - \mathbb{E}[Y_{i0}(1)|D_{i0} = 1, ARs] \leqslant 0$, proving the desired claim.

Second, we show that whether IRs or ARs have a higher expected win rate in the post period reveals the sign of $ATT^{IR} - ATT^{AR}$.

Suppose the win rate of IRs exceeds that of ARs in the post period, i.e., $\mathbb{E}[Y_{i0}(1)|D_{i0} = 1, IRs] - \mathbb{E}[Y_{i0}(1)|D_{i0} = 1, ARs] \geqslant 0$. Then the MLRP property (see the proof of Proposition 1) implies that the (potential) win rate of IRs exceeds that of ARs in the pre period as well, $\mathbb{E}[Y_{i-1}(0)|D_{i0} = 1, IRs] - \mathbb{E}[Y_{i-1}(0)|D_{i0} = 1, ARs] \geqslant 0$. We wish to show that $ATT^{IR} - ATT^{AR} \geqslant 0$, which is equivalent to showing that Equation 10 holds. But note that $\mathbb{E}[Y_{i-1}(0)|D_{i0} = 1, IRs] - \mathbb{E}[Y_{i-1}(0)|D_{i0} = 1, ARs] \geqslant 0$ and the no extreme baseline assumption together imply Equation 10 holds, as desired.

Similarly, suppose the win rate of ARs exceeds that of IRs in the post period, i.e., $\mathbb{E}[Y_{i0}(1)|D_{i0} = 1, IRs] - \mathbb{E}[Y_{i0}(1)|D_{i0} = 1, ARs] \leqslant 0$. Then $\mathbb{E}[Y_{i-1}(0)|D_{i0} = 1, IRs] - \mathbb{E}[Y_{i-1}(0)|D_{i0} = 1, ARs] \leqslant 0$. We wish to show that $ATT^{IR} - ATT^{AR} \leqslant 0$, which is equivalent to showing that Equation 11 holds. But note that $\mathbb{E}[Y_{i-1}(0)|D_{i0} = 1, IRs] - \mathbb{E}[Y_{i-1}(0)|D_{i0} = 1, ARs] \leqslant 0$ and the no extreme baseline assumption together imply Equation 11 holds, as desired.

□

B Additional Tables and Figures

In this section, we summarize the additional analyses conducted as referenced in the body of the paper. All relevant tables and figures for Appendix sections B.1-B.5 are presented at the end of Appendix B.

B.1 Data from FOIA Requests

In both Section 3 of the main body and Appendix E, we describe the process of obtaining discrimination complaint data from US states. Here, we present a collection of tables and figures that visualize this data and provide clarification on how complaints were categorized as sexual harassment versus control complaints.

Table A1 summarizes data from states that provided a substantial amount of relevant discrimination complaints that met sample restrictions for our main analysis (i.e., employment discrimination only). For each of these states, we can observe what percentage of complaints provided were classified as sexual harassment, as well as other statistics—such as the win rate across supplied complaints—calculated from the data when possible.

Figure A1, reproduced from Raj et al. (2024), provides many examples of sexual harassment that are consistent with the EEOC's definition, as discussed in Section 2.2 of the main body, and would therefore be classified as sexual harassment cases in our analysis. More serious instances, such as the example in the Sexual Assault category, may be charged as criminal offenses and, therefore, would not be included in our dataset. More information on this can be found in Section 2.3 of the main body.

Figure A2 plots the number of sexual complaints (top) and other complaints (bottom) filed over time. Note the much larger increase in sexual harassment complaints between 2017 (the year that

#MeToo began) and 2019, compared to other complaints. This suggests that the increase in sexual harassment complaints was driven by the movement and was not just reflective of a broader trend.

B.2 Results for Secondary Outcomes

While a main focus of our research is on how #MeToo impacts the win rate of sexual harassment complaints, regressions on secondary outcomes—rates of settlement and going to court—provide additional insights into the overall effects of this social movement. Section 5 of the main body presents conclusions from our win rate analysis; there is little evidence of a pre-trend, #MeToo has a positive and significant impact on complainant win probability, and female complainants do not benefit more from #MeToo than their male counterparts. The following exhibits allow for similar types of conclusions to be drawn regarding settlement rates and the likelihood of going to court.

Figure A3a shows an event study graph for settlement rates, and Figure A3b shows an event study for the probability of exiting to court. Both event studies demonstrate little evidence of a pre-trend. However, unlike with win rate, #MeToo has a negligible effect (β^{CE} = 0.001) on settlement likelihood. The movement has a slight positive effect on the probability of proceeding to court (β^{CE} = 0.007), with a transient spike at time $t = 0$. However, the post-treatment error bars indicate that this effect may not be statistically significant.

Table A2 presents estimates of equations 1a and 2a on settlement rates and the likelihood of exiting to court. As stated previously, #MeToo has essentially no effect on complaint settlement probability (β^{CE} = 0.001). The outcomes for men and women do not notably differ, as both coefficients are near zero (β^{CE}_{M} = 0.001 and β^{CE}_{W} = -0.002). Similarly, #MeToo has a positive but negligible impact on the probability of exiting to court (β^{CE} = 0.007), and outcomes across genders are consistent (β^{CE}_{M} = 0.002 and β^{CE}_{W} = 0.009).

B.3 State heterogeneity

Prior literature has found that the positive effects of the #MeToo movement for victims are concentrated in politically liberal places (Levy and Mattsson, 2025). In this appendix section, we examine the heterogeneity of our main results by state. We proceed by bisecting the states in our analysis based on their voting history in recent presidential elections (see Table A1 for details of the classification). We run our event study specification in Equation 1b separately, estimating both $\hat{\beta}^{CE}$ and $\hat{\beta}^{O}$.

We find that our results are driven entirely by blue (i.e., liberal) states, and that the combined effect of #MeToo in red and purple states is virtually zero. Figure A4 presents these event studies for blue states (top panel, $\hat{\beta}^{CE} = 0.112$), and for red and purple states (bottom panel, $\hat{\beta}^{CE} = -.004$). This pattern holds in the overlap sample as well, with overlap effects virtually zero in red and purple states but similar in magnitude and significance to our main results in the blue states. Figure A5 presents the overlap effect across these two groups. $\hat{\beta}^{O} = 0.097$ in blue states and $\hat{\beta}^{O} = -.007$

in red and purple states, compared to $\hat{\beta}^O = 0.087$ in our full sample, again highlighting that our results are being driven by blue states. Our findings are nearly unchanged if we instead group purple states with blue states.

B.4 Seasonality and Filing Invariance Assumption

Two assumptions are made in order to effectively decompose β^{CE} and draw conclusions regarding selection into reporting and changes to adjudicators' evidentiary threshold. Both of these assumptions are discussed in Section 7 of the main body. Here, we focus on the filing invariance assumption, which implies that all always reporter cases have the same treatment effects, and present a number of exhibits that boost its credibility.

The filing invariance assumption requires treatment effects for always reporters to remain stable over time. We use Figure A6 to confirm that this holds. The figure shows the win rate for overlap cases, which are used to represent always reporter cases in this paper, for the six months immediately following the start of #MeToo. The steady lines indicate that the win rate remained relatively constant.

One threat to the validity of the filing invariance assumption is seasonality in the types of cases filed, which could lead to cases filed before but resolved after #MeToo (overlap cases) being markedly different from those both filed and resolved after. Figure A7 shows that overlap cases are less likely to be filed in the winter months, indicating that there is seasonality in the volume of cases filed. The next few exhibits will prove that this volume-based seasonality does not correspond to major differences in case types or outcomes.

Table A3 lists the mean win rate for overlap cases by season, and it shows that the average win rate for winter overlap cases does not considerably differ from the other seasons. This confirms that the time of case filing, and therefore the observed volume-based seasonality, does not greatly impact treatment effects for overlap cases.

Table A4 provides additional support for the conclusion we have just drawn from Table A3. This table presents the overlap results for each individual season, and it demonstrates that the magnitude of treatment is largely similar for all cases—regardless of when they were filed.

Table A5 compares β^{CE} to β^O by filing season for complaints with gender, but it also allows for an important comparison of the seasonal estimates to the corresponding values from our main results. For both the β^O and β^{CE} estimates, the seasonal TWFE values are all relatively similar to the main results. All seasonal β^O values are within 4.2 pp of the main β^O = 0.102, and the seasonal β^{CE} values are within 7 pp of the main β^{CE} = 0.116, with the winter and summer estimates being within 0.8 pp. This strengthens the A2 assumption by confirming that the time of filing does not greatly impact the relative magnitudes of beta estimates, and thus that treatment effects remain consistent across seasons. Beyond the size of the estimates, 75% of the seasonal inequalities have the same sign as their corresponding main results, with inconsistencies concentrated in the fall.

Because the pre-period includes only one month in the fall of 2017, a disproportionate number of that season's overlap cases were filed a full year before #MeToo, making it likely that fall cases are qualitatively different than those in other seasons.

B.5 Secondary Outcomes for Overlap Cases

In Section 7.1 of the main body, we present the results of estimating equations 1a and 2a on win rate for the overlap sample. Here, we present overlap results for secondary outcomes.

Table A6 presents estimates of Equations 1a and 2a on settlement rates and the likelihood of exiting to court for overlap cases. #MeToo has a minimal negative effect on overlap case settlement (β^O = -0.022), with comparable effects across genders (β^O_M = -0.023 and β^O_W = -0.017). Similarly, #MeToo has a very small, negative impact on the probability of exiting to court (β^O = -0.013), and outcomes remain consistent across genders (β^O_M = -0.019 and β^O_W = -0.012). β^O is slightly smaller than β^{CE} for both the settlement ($\beta^O = -0.022 < \beta^{CE} = 0.001$) and court outcomes ($\beta^O = -0.013 < \beta^{CE} = 0.007$), suggesting that induced reporters are positively selected in both cases but that the extent of this selection is minimal.

B.6 Overlap and Always Reporter Cases

In Appendix B.4, as well as in Section 7.1 of the main body, we mention that overlap cases are used as a proxy for always reporter cases. Here, we demonstrate that this is reasonable by verifying that overlap cases are not characteristically different from other always reporter cases.

Figure A8 plots the average duration of cases that are open in a given month in the year preceding #MeToo, for both sexual harassment and non-sexual harassment cases. The trend is relatively flat month-to-month suggesting that the distribution of case durations is relatively stable before #MeToo. This evidence lends credence to our claim in Section 7.1 that our overlap analysis is unlikely to be be affected by selected on duration in a way that biases our results.

Table A7 compares actual overlap cases to placebo cases, or those that would have been classified as overlap if #MeToo had begun in 2016, based on six (exogenous-to-adjudicators) characteristics. The differences between the actual and placebo cases, while statistically significant in many outcomes, are likely economically small, lending support to overlap cases being broadly representative of always reporter cases.

B.7 Regression Discontinuity

In Section 8, we discuss how directly comparing win rates for cases filed (just) before and (just) after #MeToo may misrepresent the direction or magnitude of the selection effect. As previously discussed, our difference-in-difference analysis, using data from the entire pre- and post-periods, finds that female induced reporters are positively selected for foundedness while male induced reporters are negatively selected. In this section, we present regression discontinuity (RD) analyses

that suggest that the "immediate" selection effect directionally matches the overall selection effect for both genders.

In each RD design, we restrict to sexual harassment cases resolved after #MeToo (but filed either before or after). The running variable is the time of case filing. Therefore, the RD effect is the difference in win rate between always reporters (left of the discontinuity) and always *plus* induced reporters (right of the discontinuity). Recalling Remark 1, the sign of the RD coefficient is the same as the sign of the RD coefficient if we could restrict to only induced reporters in the post-period. A positive RD effect is therefore consistent with $ATT^{IR} > ATT^{AR}$, and vice versa. Recalling that we do not observe the share of induced reporters in the post period $(1 - \omega)$, the magnitude of these effects should not be literally interpreted, and indeed, are mechanically a lower bound on the actual effects—our effects are unbiased if and only if $\omega = 0$.

In Table A8, we present estimates of how the win rate changes around the #MeToo discontinuity for men, and for women. Each row represents a different specification, varying the fixed effects and the polynomial order. All estimates for men are negative—consistent with male IRs being negatively selected—and all estimates for women are positive (consistent with female IRs being positively selected). In Figure A9, we plot our preferred RD specification using state by time fixed effects and polynomial of order two, to visually confirm the direction of the point estimates.

TABLE A1 — DATA SUMMARY BY STATE FOR EMPLOYMENT DISCRIMINATION COMPLAINTS

State	N	Sexual harassment	Complainant won	$ paid to complainant	Female complainant
CA	90,220	.247	.000	.	.
NY	48,183	.144	.270	.	.491
MA	17,540	.073	.116	.	.522
WI	16,135	.064	.	.	.
MI	7,350	.043	.0004	$9,155	.
FL	5,110	.005	.050	.	.
AK	2,028	.133	.067	$10,465	.
KY	1,862	.028	.007	$15,400	.
ND	1,801	.165	.024	$1,656	.
HI	1,756	.100	.002	$11,767	.
N	191,985	191,978	74,029	1,060	64,499

Table A1: This table presents means of variables describing employment discrimination complaints, for each of the 10 states that provided such data (see Appendix E for more details on our FOIA request process). N indicates the number of non-missing observations. Dots indicate missing data that states either do not record or do not disclose via FOIA. Sexual harassment is the share of sexual harassment complaints out of all discrimination cases filed. "Complainant won" is the win rate across all cases in that state. The win rate in California is 0% because California mostly adjudicates complainants in formal court procedures (only 3.4% of complainants ever win at the investigation stage, and this number is 0% after meeting our sample restrictions). We used a Python API to determine the gender of the complainant based on their name. States are colored by voting history in the 2012, 2016, 2020, and 2024 presidential elections, with red (blue) states indicating that Republicans (Democrats) won that state in the majority of the last four elections, and purple indicating each party won two elections, as taken from https://en.wikipedia.org/wiki/Red_states_and_blue_states.

FIGURE A1 — DEFINITIONS OF SEXUAL WRONGDOING

VERBAL SEXUAL HARASSMENT

- Someone whistling, honking, making kissy noises, "Pssst" sounds, or leering/staring aggressively at you.
- Someone saying things like, "Hey Baby," "Mmmm Sexy," "Yo Shorty," "Mami/Mamacita," "Give me a smile," or similar comments in a way that is disrespectful and/or unwanted and/or makes you feel unsafe.
- Someone calling you a sexist slur, like "Bitch," "Slut," "Cunt," or "Whore".
- Someone purposefully misgendering you or calling you a homophobic or transphobic slur, like "Fag," "Dyke," or "Tranny."
- Someone talking about your body parts inappropriately or offensively (such as your legs, crotch, butt, or breasts), saying sexually explicit comments ("I want to do BLANK to you") or asking inappropriate sexual questions.

CYBER SEXUAL HARASSMENT

- Someone sending you sexual content without your permission, sexualizing you inappropriately, or threatening you with sex acts, in *private online spaces*, such as over e-mail or text.
- Someone sending you sexual content without your permission, sexualizing you inappropriately, or threatening you with sex acts, in *public online spaces*, such as on social media.
- Someone taking and/or sharing sexual pictures or videos of you without your permission.

DATING COERCION OR SEXUAL THREATS

- Someone making threats to harm you, to harm someone you know, or to share personal information you don't want shared (examples could include your sexual orientation or drug use history or immigration status).
- Someone saying you must date them or do a sexual act for them in exchange for something (such as a good grade, a promotion, a job, drugs, food, or something similar) or instead of something (like paying rent or a citation, etc).
- Someone repeatedly texting or calling you in a harassing way.
- Someone repeatedly asking you for a date or your phone number when you've said no or ignored them.

PHYSICALLY AGGRESSIVE SEXUAL HARASSMENT

- Someone flashing or exposing their genitals to you without your permission.
- Someone physically following you without your permission.
- Someone purposely touching you or brushing up against you in an unwelcome, sexual way.

SEXUAL ASSAULT

- Someone forcing you to do a sexual act without your permission or one that you didn't want to do (including while you were under the influence of alcohol or drugs).

Figure A1: Definitions of different types of sexual wrongdoing, taken from Raj et al. (2024).

FIGURE A2 — NUMBER OF COMPLAINTS FILED OVER TIME

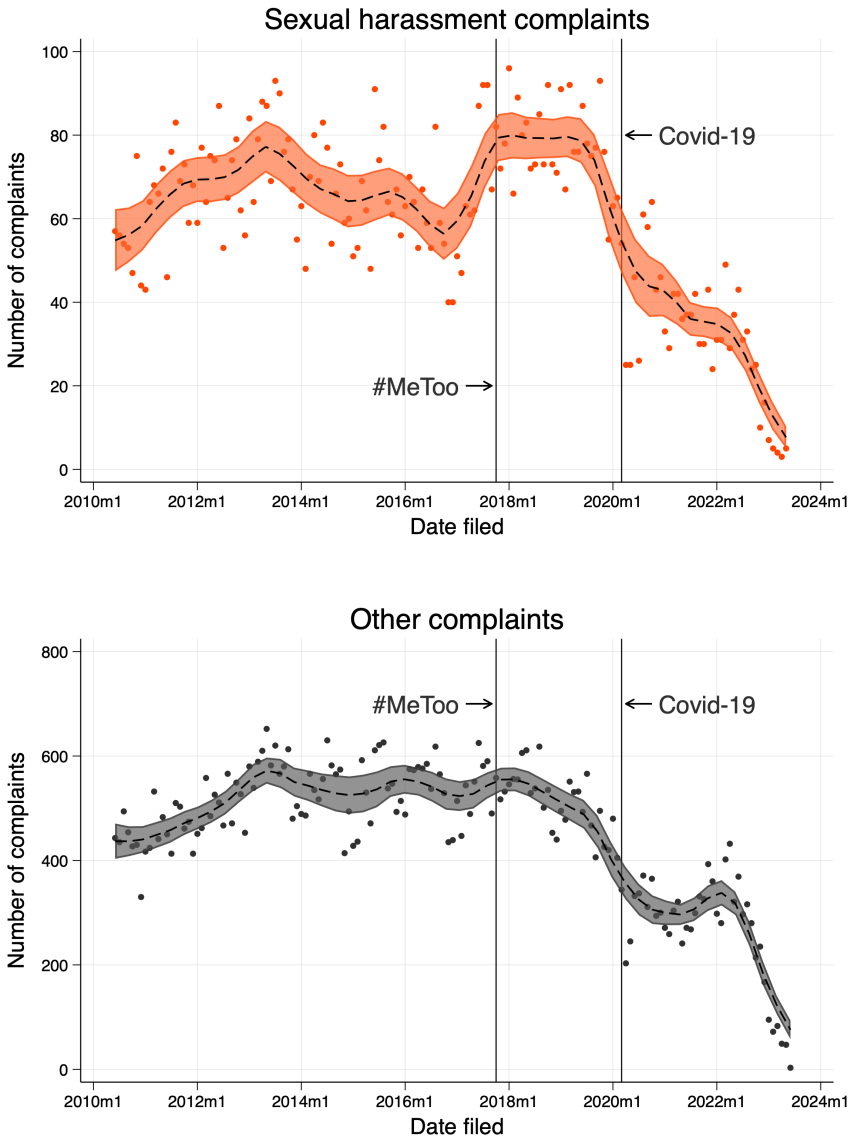


Figure A2: The figure shows the number of employment discrimination complaints filed over time, excluding California, Florida, and Wisconsin. These states are omitted because data for them begin only in early 2017, which would otherwise create an artificial increase in filings following the treatment period.

FIGURE A3a — IMPACT ON SETTLEMENT

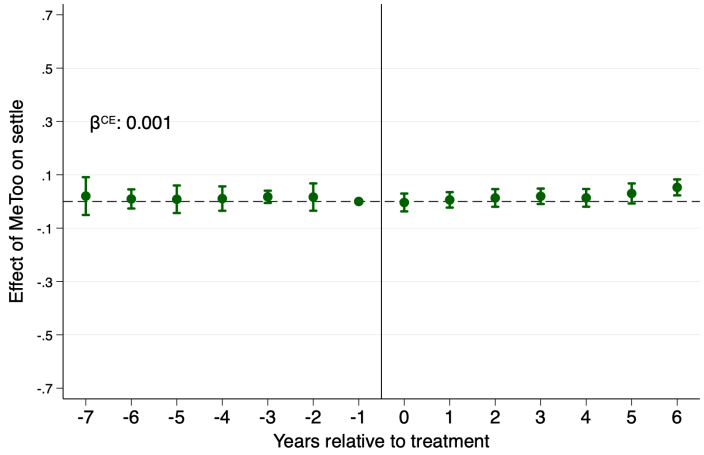


FIGURE A3b — IMPACT ON EXITING TO COURT

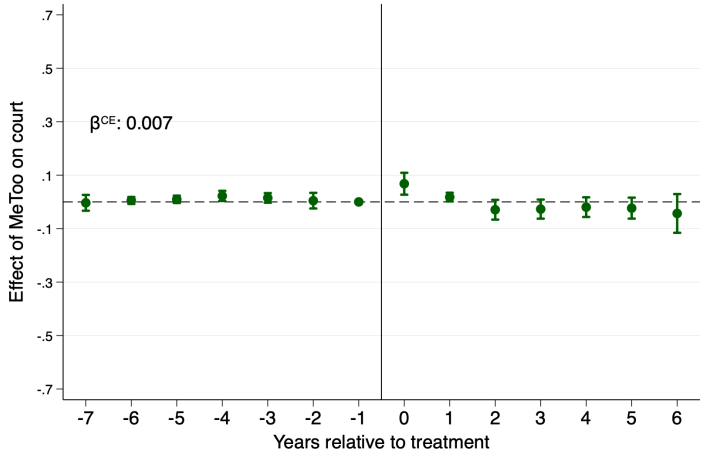


Figure A3: These figures provide estimates of Equation 1b. The x-axis in all subfigures represents 12-month intervals centered around the #MeToo start date. $t = -1$ is the 12 months preceding October 15, 2017, $t = 0$ is the 12-month period following October 15, 2017, and so on. We control for unit-by-state and year-month-by-state fixed effects, where the "year-month" specifies the date of the complaint resolution. 95% confidence intervals are reported using standard errors clustered at the complaint type by state level.

TABLE A2 — EFFECTS OF #METOO ON ALL COMPLAINT OUTCOMES

	Settled			Court		
	All complaints	**Complaints with gender**		**All complaints**	**Complaints with gender**	
SH × Post	0.001	-0.001	0.001	0.007	0.009	0.002
	(0.004)	(0.003)	(0.009)	(0.006)	(0.007)	(0.012)
SH × Post × Female			-0.003			0.007
			(0.009)			(0.009)
Unit and Time × State FE	✓	✓	✓	✓	✓	✓
Unit and Time × State × Female FE			✓			✓
N	175,495	64,499	64,499	175,801	64,499	64,499
R^2	0.099	0.103	0.110	0.631	0.140	0.150
Control mean	0.071	0.094	0.094	0.388	0.048	0.048

Table A2: Settled is whether a case was settled. Court is whether a case was taken to court. Time fixed effects are year-month of case resolution. Standard errors clustered at the case type level. Single asterisk denotes statistical significance at the 90% level of confidence; double, 95%; triple, 99%.

FIGURE A4 — HETEROGENEITY OF AGGREGATE RESULTS BY STATE TYPE

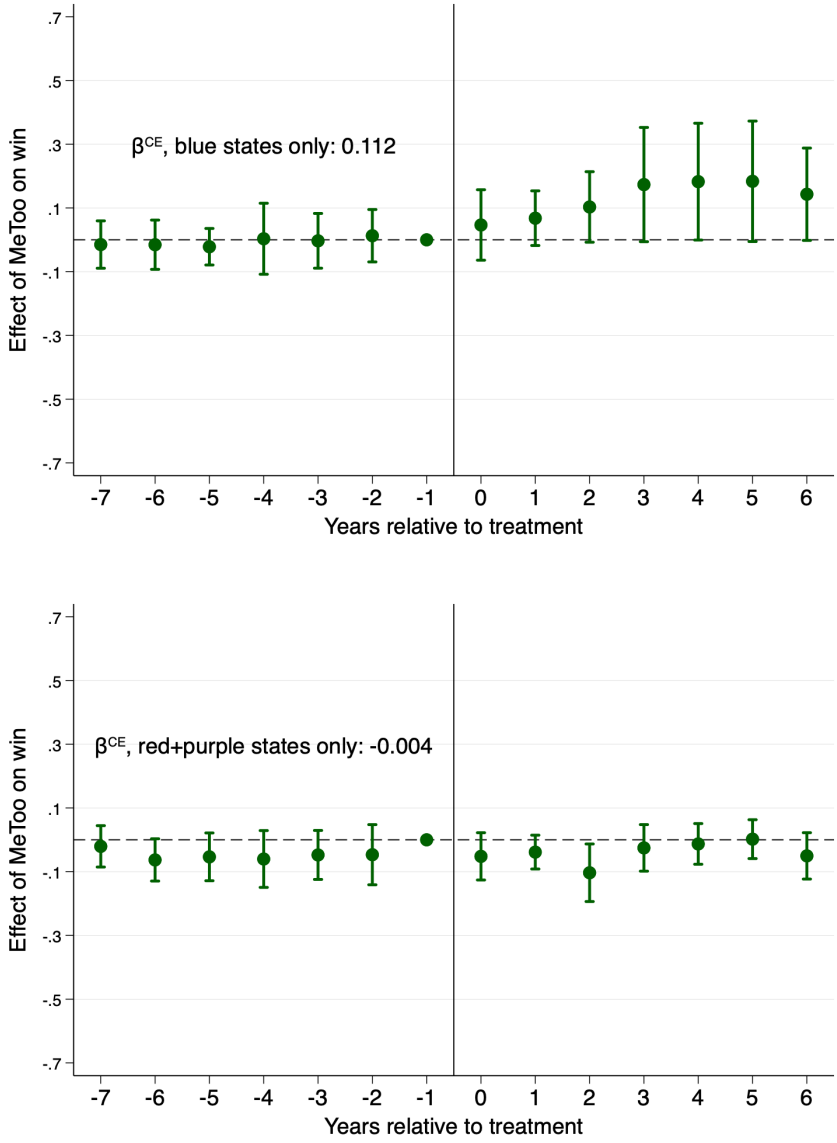


Figure A4: The figure plots the results of Equation 1b estimated on the group blue states (top panel), and red and purple states combined (bottom panel). Table A1 describes our classification of states into red, blue, and purple.

FIGURE A5 — HETEROGENEITY OF OVERLAP RESULTS BY STATE TYPE

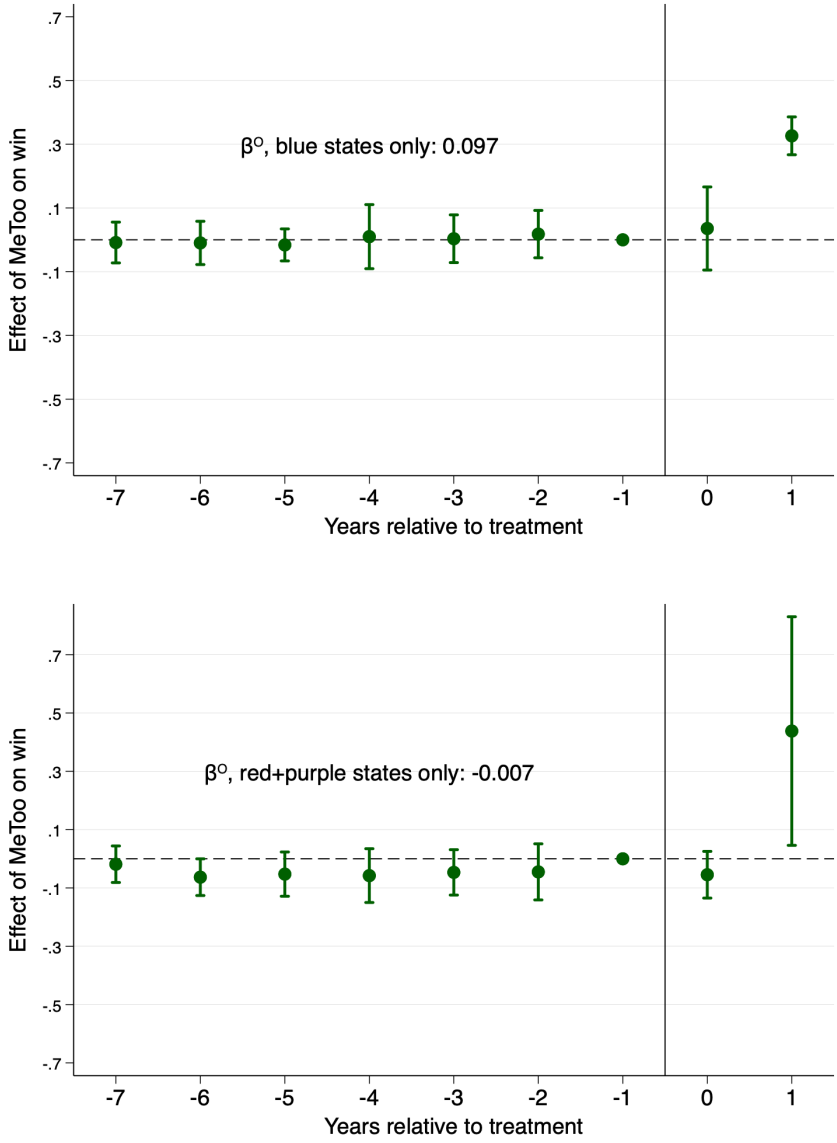


Figure A5: The figure plots the results of Equation 1b estimated on the group blue states (top panel), and red and purple states combined (bottom panel) for overlap complaints. Table A1 describes our classification of states into red, blue, and purple.

FIGURE A6 — WIN RATE OF OVERLAP CASES RESOLVED AFTER #METOO

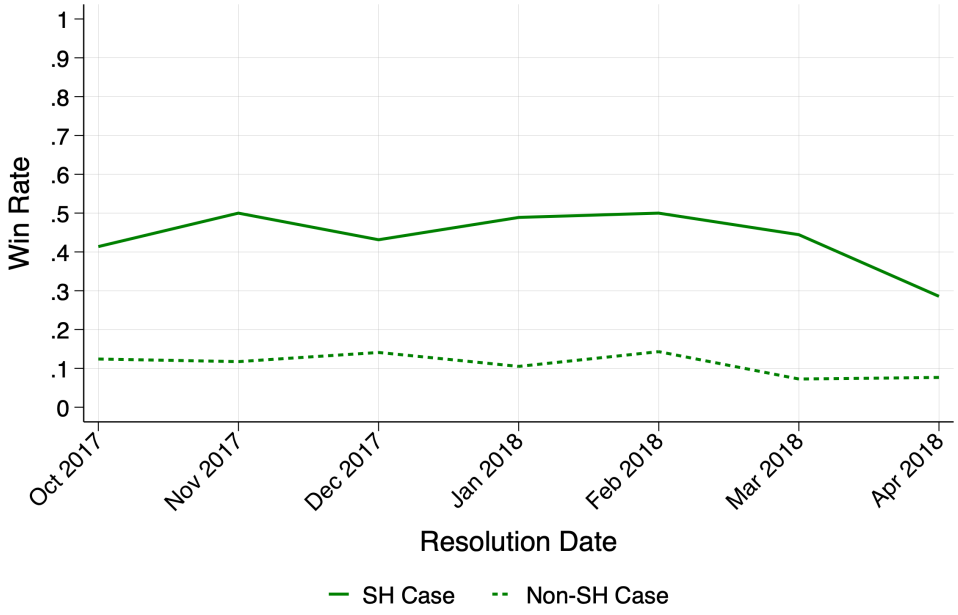


Figure A6: This figure shows the win rate of both sexual harassment and control (non-sexual harassment) cases over time, for cases resolved in the six months post-#MeToo.

FIGURE A7 — PERCENT OF OVERLAP CASES FILED IN EACH CALENDAR MONTH

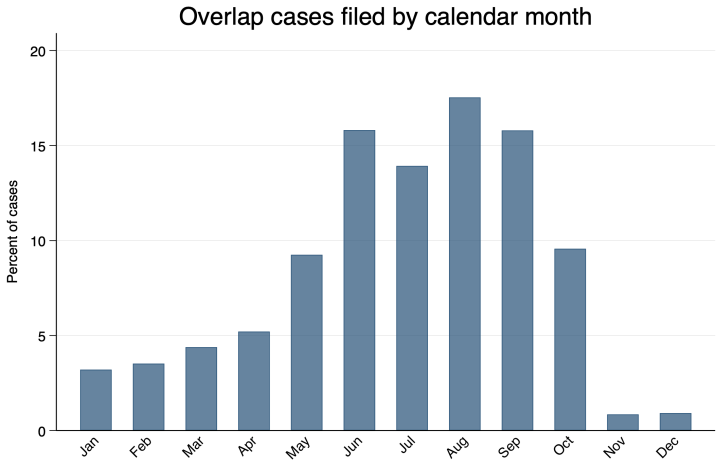


Figure A7: This figure displays the percent of overlap cases (both control and treatment) that are filed in each month.

TABLE A3 — COMPLAINANT WIN RATE AND CASE DURATION BY SEASON (OVERLAP CASES)

Season	Win Rate (Mean)	Duration (Mean)
Spring	0.1008	354.99
	(0.3011)	(158.41)
Summer	0.1633	297.07
	(0.3697)	(166.79)
Fall	0.1252	303.85
	(0.3311)	(181.80)
Winter	0.0573	443.52
	(0.2327)	(141.65)
Total	0.1268	324.79
	(0.3328)	(172.76)

Table A3: This table displays the average win rate (Col 1) and duration (Col 2) of overlap cases filed in each season. Duration is measured in days and refers to the amount of time between the filing and resolution dates of a case. Standard deviations are in parentheses.

TABLE A4 — MAIN OVERLAP RESULTS BY SEASON OF CASE FILING

	All complaints	Complaints with gender	
Panel A: Winter			
SH × Post	0.054	0.069	0.230
	(0.068)	(0.081)	(0.389)
SH × Post × Female			-0.179
			(0.411)
Unit and Time × State FE	✓	✓	✓
N	8,365	6,567	6,543
R^2	0.175	0.126	0.149
Control mean	0.152	0.226	0.229
Panel B: Spring			
SH × Post	0.123*	0.141*	0.275
	(0.066)	(0.071)	(0.163)
SH × Post × Female			-0.223
			(0.134)
Unit and Time × State FE	✓	✓	✓
N	9,807	7,644	7,626
R^2	0.169	0.116	0.134
Control mean	0.152	0.226	0.229
Panel C: Summer			
SH × Post	0.052***	0.060**	0.001
	(0.019)	(0.023)	(0.059)
SH × Post × Female			0.106
			(0.066)
Unit and Time × State FE	✓	✓	✓
N	11,232	8,948	8,933
R^2	0.171	0.118	0.140
Control mean	0.152	0.226	0.229
Panel D: Fall			
SH × Post	0.133***	0.136***	0.275***
	(0.042)	(0.042)	(0.015)
SH × Post × Female			-0.185***
			(0.045)
Unit and Time × State FE	✓	✓	✓
Unit and Time × State × Female FE			✓
N	9,566	7,550	7,526
R^2	0.188	0.139	0.160
Control mean	0.152	0.226	0.229

Table A4: Each panel is equivalent to Table 3, except that the sample is restricted to overlap cases filed within the given season.

TABLE A5 — COMPARISON OF $\hat{\beta}^{CE}$ AND $\hat{\beta}^{O}$ BY FILING SEASON FOR COMPLAINTS WITH GENDER

Season Filed	$\hat{\beta}^{CE}$	$\lessgtr$	$\hat{\beta}^{O}$
All			
TWFE	**0.116**	$>$	**0.102**
Triple Diff (Men)	0.119	$<$	0.153
Triple Diff (Women)	0.123	$>$	0.086
Winter			
TWFE	0.120	$>$	0.069
Triple Diff (Men)	0.083	$<$	0.230
Triple Diff (Women)	0.132	$>$	0.051
Spring			
TWFE	0.158	$>$	0.141
Triple Diff (Men)	0.199	$<$	0.275
Triple Diff (Women)	0.129	$>$	0.052
Summer			
TWFE	0.108	$>$	0.060
Triple Diff (Men)	0.048	$>$	0.001
Triple Diff (Women)	0.164	$>$	0.107
Fall			
TWFE	0.046	$<$	0.136
Triple Diff (Men)	0.101	$<$	0.275
Triple Diff (Women)	0.027	$<$	0.090

Table A5: Each cell represents the coefficient from estimating Equation 1a and 2a on the full sample (Col 1) and overlap sample (Col 2) while restricting to cases filed in each season.

FIGURE A8 — AVERAGE DURATION OF CASES OPEN IN EACH MONTH BEFORE #METOO

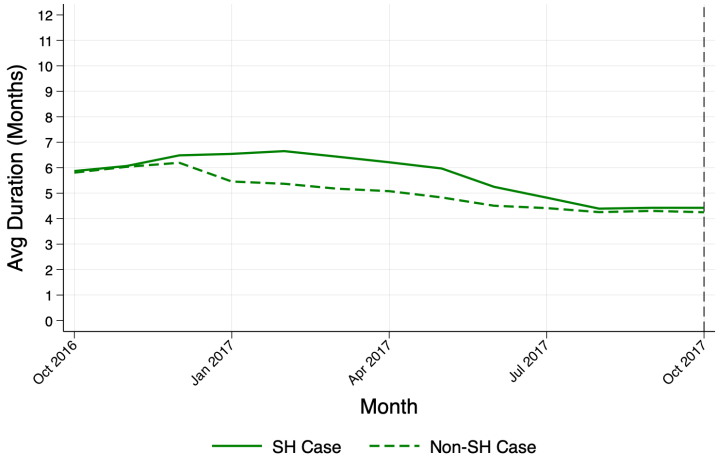


Figure A8: This figure displays the average duration of all open cases in a given month for the 12 months before #MeToo (indicated with grey dashed line) by sexual harassment and non-sexual harassment cases.

TABLE A6 — EFFECTS OF #METOO ON ALL COMPLAINT OUTCOMES, OVERLAP CASES

	Settled			Court		
	All complaints	**Complaints with gender**		**All complaints**	**Complaints with gender**	
SH × Post	-0.022***	-0.021***	-0.023*	-0.013	-0.012	-0.019*
	(0.006)	(0.006)	(0.011)	(0.018)	(0.020)	(0.009)
SH × Post × Female			0.006			0.007
			(0.013)			(0.017)
Unit and Time × State FE	✓	✓	✓	✓	✓	✓
Unit and Time × State × Female FE			✓			✓
N	59,471	41,628	41,627	59,488	41,628	41,627
R^2	0.109	0.108	0.115	0.656	0.129	0.138
Control mean	0.071	0.094	0.094	0.388	0.048	0.048

Table A6: Settled is whether a case was settled. Court is whether a case was taken to court. Time fixed effects are year-month of case resolution. Standard errors clustered at the case type level. Single asterisk denotes statistical significance at the 90% level of confidence; double, 95%; triple, 99%.

TABLE A7 — COMPARISON OF ACTUAL OVERLAP CASES (2017) VERSUS PLACEBO OVERLAP CASES (2016)

Variable	(1) 2016	(2) 2017	(3) Diff	(4) p-value
Complainant is female	0.51	0.50	-0.01	0.28
Sexual harassment	0.10	0.08	-0.02	0.00***
Filed in Spring	0.24	0.21	-0.03	0.00***
Filed in Summer	0.39	0.44	0.05	0.00***
Filed in Fall	0.25	0.25	-0.00	1.00
Filed in Winter	0.12	0.10	-0.02	0.00***
Observations	5,315	6,230	11,545	

Table A7: From top to bottom, the variables measure the share of cases with a female complainant, the share classified as sexual harassment, and the share filed in each season. Col 1 presents these values for cases that would have been considered overlap if #MeToo had occurred in 2016, while Col 2 presents values for cases that are actually considered overlap. Col 3 shows the difference in values for each variable (Col 1 minus Col 2), and Col 4 displays the corresponding p-values. Single asterisk denotes statistical significance at the 90% level of confidence; double, 95%; triple, 99%.

TABLE A8 — RD POINT ESTIMATES ACROSS VARIOUS SPECIFICATIONS

	Polynomial order	RD effect	p-value
Men			
No fixed effects	1	−0.180	0.300
No fixed effects	2	−0.275	0.431
State and time fixed effects	1	−0.137	0.460
State and time fixed effects	2	−0.199	0.787
State by time fixed effects	1	−0.120	0.498
State by time fixed effects	2	−0.190	0.871
Women			
No fixed effects	1	0.024	0.740
No fixed effects	2	0.048	0.841
State and time fixed effects	1	0.051	0.500
State and time fixed effects	2	0.096	0.618
State by time fixed effects	1	0.025	0.547
State by time fixed effects	2	0.085	0.649

Table A8: This table presents estimates of RD effects and associated p-values analogous to Figure A9, but across six different specifications, including the one shown in Figure A9. We first restrict the sample to sexual harassment complaints resolved after #MeToo filed by male (top panel) and female (bottom panel) complainants. Each row presents the result from a regression discontinuity design where the outcome is the win rate, and the specification changes row-by-row as indicated. We use an optimal bandwidth selection procedure (Calonico et al., 2014) to select 200 days for men and 240 days for women as the optimal bandwidth.

FIGURE A9 — RD PLOT OF WIN RATE FOR CASES RESOLVING AFTER #METOO

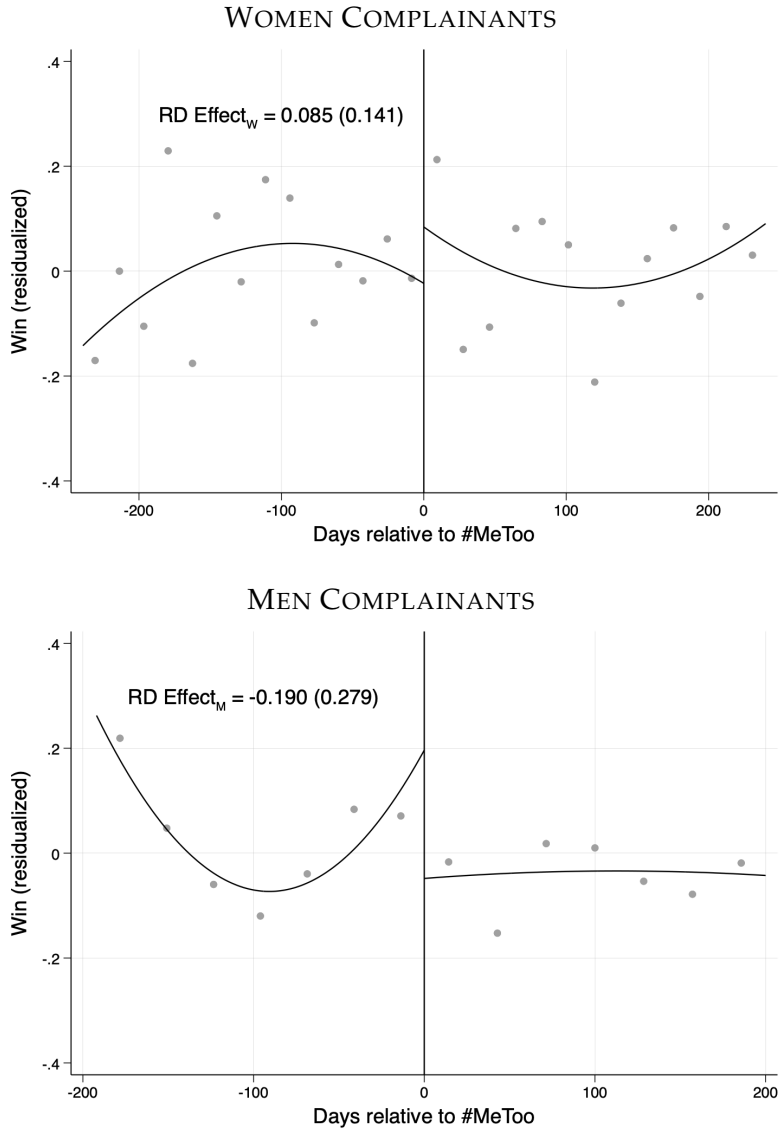


Figure A9: These figures estimate a regression discontinuity design using an optimal bandwidth selection method (Calonico et al., 2014). Each figure shows the win rate change around the #MeToo discontinuity for sexual harassment cases that resolve after #MeToo, filed by women (top) versus men (bottom). The x-axis represents the day each case for filed, therefore, data points before #MeToo represent win rates for cases filed before but resolved after #MeToo, data points shown after #MeToo represent win rates for cases filed after and resolved after. The chosen bandwidth was 240 days around #MeToo for women and 200 for men. Therefore, the x-axis in the top panel represents 240-day intervals centered around the #MeToo start date, and bottom represents 200-day intervals. The outcome is win rate, residualized on state fixed effects interacted with time fixed effects (representing the year-month of complaint resolution) with a polynomial fit of order two. The "RD Effect" label in each graph shows the point estimate and SE in parentheses, corresponding to the specification in the last row in each panel of Table A8.

C Robustness

In this section, we present a variety of robustness checks to the main results presented in Section 5. Our goal is to investigate our main empirical findings in alternative samples and under alternative definitions of control complaints. For each robustness check we consider, we investigate two sets of findings. First, whether both β_M^O and β_W^O are positive, which implies by Propositions 1 and 2 that the evidentiary threshold adjudicators used when evaluating complaints lowered for sexual harassment cases following #MeToo for both genders. Second, whether male or female induced reporters were positively or negatively selected on the dimension of foundedness compared to always reporters, which we investigate by comparing the ordinal rankings of the combined effect coefficients (β_M^{CE} and β_W^{CE}, respectively) to the overlap coefficients (β_M^O and β_W^O, respectively) (again, see Propositions 1 and 2).

As a summary of our findings, we first present Table A9, which provides a comparison of the coefficients for each of our alternative samples and definitions of control complaints. Broadly, these results are similar to those presented in the main analysis. We describe each of these robustness checks below.

In the first robustness check we consider, we restrict the sample to complaints filed under only one basis of discrimination. This removes ambiguity from multi-basis cases, which may involve overlapping types of discrimination.[30] The estimated effect on the probability of winning remains positive and significant for both combined and overlap coefficients. $\hat{\beta}_M^{CE}$ is smaller than $\hat{\beta}_M^O$, and $\hat{\beta}_W^{CE}$ is larger than $\hat{\beta}_W^O$, both of which directionally align with our main results. This implies that male reporters are negatively selected and female reporters are positively selected, even for cases filed under one basis of discrimination, which is consistent with our main results. The full results for this robustness check are presented in Tables A10 and A11.

In the second robustness check we consider, we exclude complaints alleging retaliation from the analysis. In our main specification, we retain retaliation complaints as part of the control group, but retaliation cases may blur the distinction between treatment and control if they are responses to earlier sexual harassment. Excluding these complaints yields nearly identical results to our main specification for both men and women. Both $\hat{\beta}_M^O$ and $\hat{\beta}_W^O$ remain positive and significant. The direction of selection is also consistent with our main results for both men and women. The combined effect for men $\hat{\beta}_M^{CE}$ (0.113) is smaller than $\hat{\beta}_M^O$ (0.153), and the combined effect for women $\hat{\beta}_W^{CE}$ (0.119) is larger than $\hat{\beta}_W^O$ (0.082), both of which are directionally consistent with our main results. The estimates for this robustness check are presented in Tables A10 and A11.

In the third robustness check we consider, we restrict our sample to complaints filed by the start of the Covid-19 pandemic in March 2020. This also addresses concerns that post-2020 complaints

[30] It is difficult to determine using our data how single- versus multi-tagged complaints differ on unobservables. On observables, complainants of single-tagged cases are as likely to win their case as their multi-tagged counterparts (15.6% versus 15.9%), but single-tagged cases last longer (241 days versus 105 days), are more likely to be dismissed or settled, and are less likely to exit to court than multi-tagged cases.

are systematically different as a result of pandemic-related changes in workplace dynamics or agency operations. Because the start of the Covid-19 pandemic occurs after the start of #MeToo, $\hat{\beta}^{O}$, $\hat{\beta}^{O}_{M}$, and $\hat{\beta}^{O}_{W}$ are mechanically unchanged after this sample restriction. However, $\hat{\beta}^{CE}$, $\hat{\beta}^{CE}_{M}$, and $\hat{\beta}^{CE}_{W}$ are potentially affected by the restriction, and we compare the differences between the overlap and combined effects, respectively. $\hat{\beta}^{O}_{M}$ is larger than $\hat{\beta}^{CE}_{M}$ for men who filed before Covid-19 ($0.153 > 0.057$), which also directionally aligns with the main results in indicating that men are negatively selected. We find that $\hat{\beta}^{O}_{W}$ is also larger than $\hat{\beta}^{CE}_{W}$ for women who filed before Covid-19 ($0.086 > 0.074$), which suggests that women 'early filers' are negatively selected. Since our main results show that women's complaints are stronger after #MeToo, this suggests post-Covid-19 women's cases are potentially driving the positive selection. The estimates for this robustness check are presented in Tables A10 and A11.

In the fourth robustness check we consider, we present estimates of Equation 1a using case filing rather than case resolution time fixed effects. Throughout the paper, we have defined time fixed effects at case resolution in order to capture the effect at the time of adjudicator decision. However, since overlap cases are filed and resolved on either side of treatment, the choice of time fixed effects may be consequential for the ordinal comparison of $\hat{\beta}^{CE}$ and $\hat{\beta}^{O}$. Table A12 shows that using filing fixed effects changes the point estimates of $\hat{\beta}^{CE}$, $\hat{\beta}^{O}$, and their gendered version only slightly, significance of all point estimates remains the same, and the main conclusions of the paper are unchanged to using case filing fixed effects.

In the fifth robustness check, we examine how sensitive our inference is to our choice of clustered standard errors. Throughout the paper we cluster standard errors at the state-by-discrimination type level to account for within-state and within-discrimination type correlation in how complaints are filed and adjudicated, yielding 57 clusters. A natural robustness check is wild bootstrap – however, since treatment is assigned at the discrimination type level $g(i)$ with a single treated discrimination type, cluster-robust inference at this level would be unreliable even with wild bootstrap. We therefore conduct a permutation test by reassigning treatment to each of the other six (non-treated) discrimination types and re-run the regression described in Equation 1a. If the effect of treatment on sexual harassment is truly an outlier, we should expect small (in magnitude) and statistically insignificant coefficients on these placebo treatments. Moreover, we apply a pre-trends restriction to these placebo treatments to ensure that we are not comparing a plausibly causal estimate (our baseline) to non-causal estimates driven by pre-trends in the candidate placebo treatments. Specifically, we include only placebo treatments in which no more than one pre-#MeToo treatment estimate is statistically different from 0 at the $\alpha = 0.05$ level, as estimated in Equation 1b. The following placebos are retained: age, disability, nationality, and race. The results, presented in Table A13, are encouraging – the actual treatment effect (0.101) exceeds all placebo estimates in magnitude and has a smaller p-value (0.001).

In the final robustness check we consider, Table A14 presents our main results using a broader range of discrimination complaints filed across housing, public accommodations, education, and

employment sectors. Both $\hat{\beta}^{CE}$ and $\hat{\beta}^{O}$ remain positive and significant (0.130, std. err. = 0.072 and 0.176, std. err. = 0.041), indicating that adjudicators lowered their evidentiary threshold after #MeToo even when pooling cases across multiple areas of public life. Similarly, the direction of selection for both genders is consistent with our main results. $\hat{\beta}^{O}_{M}$ is larger than $\hat{\beta}^{CE}_{M}$ ($0.207 > 0.203$), while $\hat{\beta}^{CE}_{W}$ is larger than $\hat{\beta}^{O}_{W}$ ($0.249 > 0.204$), indicating men are negatively selected while women are positively selected among all discrimination complaints in our sample. These findings are presented in Tables A14 and A15.

TABLE A9 — COMPARISON OF $\hat{\beta}^{CE}$ AND $\hat{\beta}^{O}$ ACROSS ROBUSTNESS CHECKS FOR COMPLAINTS WITH GENDER

Main results	$\hat{\beta}^{CE}$	$\lessgtr$	$\hat{\beta}^{O}$
Triple Diff (Men)	0.119	$<$	0.153
Triple Diff (Women)	0.123	$>$	0.086
Robustness Check	$\hat{\beta}^{CE}$	$\lessgtr$	$\hat{\beta}^{O}$
Single-tagged			
Triple Diff (Men)	0.084	$<$	0.144
Triple Diff (Women)	0.085	$>$	0.067
No retaliation			
Triple Diff (Men)	0.113	$<$	0.153
Triple Diff (Women)	0.119	$>$	0.082
Pre-Covid			
Triple Diff (Men)	0.057	$<$	0.153
Triple Diff (Women)	0.074	$<$	0.086
All jurisdictions			
Triple Diff (Men)	0.203	$<$	0.207
Triple Diff (Women)	0.249	$>$	0.204

Table A9: Each cell represents the coefficient from estimating Equation 1a on the full sample (Col 1) and overlap sample (Col 2) while restricting to cases filed on only one basis (Single-tagged), excluding retaliation-based cases in the control group (No retaliation), only incorporating cases filed before the start of the Covid-19 pandemic (Pre-Covid), or including all jurisdictions of discrimination.

TABLE A10 — ROBUSTNESS OF $\hat{\beta}^{CE}$ ACROSS DIFFERENT SAMPLES AND CONTROL GROUPS

	Single-tagged			No retaliation			Pre-Covid		
	All complaints	**Complaints with gender**		**All complaints**	**Complaints with gender**		**All complaints**	**Complaints with gender**	
SH × Post	0.072**	0.077*	0.084***	0.100***	0.112***	0.113***	0.053***	0.062***	0.057***
	(0.036)	(0.038)	(0.023)	(0.027)	(0.026)	(0.027)	(0.016)	(0.016)	(0.013)
SH × Post × Female			0.001			0.006			0.017
			(0.052)			(0.021)			(0.020)
Unit and Time × State FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
N	37,714	26,466	26,462	66,727	42,339	42,334	57,824	40,616	40,611
R^2	0.168	0.109	0.121	0.177	0.092	0.100	0.144	0.074	0.081
Control mean	0.153	0.209	0.210	0.156	0.231	0.234	0.165	0.226	0.228

Table A10: This table is equivalent to Table 2, but the estimation sample is limited to cases filed on only one basis, excludes retaliation-based cases in the control group, or includes only cases filed before the start of the Covid-19 pandemic (March 2020), respectively.

TABLE A11 — ROBUSTNESS OF $\hat{\beta}^O$ ACROSS DIFFERENT SAMPLES AND CONTROL GROUPS

	Single-tagged			No retaliation			Pre-Covid		
	All complaints	**Complaints with gender**		**All complaints**	**Complaints with gender**		**All complaints**	**Complaints with gender**	
SH × Post	0.073***	0.082***	0.144***	0.088***	0.101***	0.153***	0.087***	0.102***	0.153***
	(0.011)	(0.010)	(0.041)	(0.014)	(0.012)	(0.010)	(0.013)	(0.011)	(0.014)
SH × Post × Female			-0.077			-0.071**			-0.067**
			(0.055)			(0.027)			(0.027)
Unit and Time × State FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
N	23,374	18,387	18,379	35,587	27,658	27,650	39,321	30,741	30,733
R^2	0.152	0.104	0.115	0.134	0.077	0.084	0.128	0.073	0.081
Control mean	0.153	0.209	0.210	0.156	0.231	0.234	0.165	0.226	0.228

Table A11: This table is equivalent to Table 3, but the estimation sample is limited to cases filed on only one basis, excludes retaliation-based cases in the control group, or includes only cases filed before the start of the Covid-19 pandemic (March 2020), respectively.

TABLE A12 — ROBUSTNESS OF MAIN RESULTS TO USING COMPLAINT FILING FIXED EFFECTS

	Combined effect	Overlap
$\hat{\beta}$	0.100***	0.090***
	(0.029)	(0.011)
$\hat{\beta}_M$	0.118***	0.156***
	(0.025)	(0.014)
$\hat{\beta}_W$	0.120***	0.085***
	(0.033)	(0.021)

Table A12: This table reports estimates of Equation 1a for the entire sample and the overlap sample using complaint filing rather than complaint resolution time fixed effects $\mu_{m(i)s(i)}$. Standard errors are reported in parentheses. Coefficients in the table represent the combined and overlap effects for all complaints ($\hat{\beta}^{CE}$ and $\hat{\beta}^{O}$), for men ($\hat{\beta}_M^{CE}$ and $\hat{\beta}_M^{O}$), and women ($\hat{\beta}_W^{CE}$ and $\hat{\beta}_W^{O}$). Single asterisk denotes statistical significance at the 90% level of confidence; double, 95%; triple, 99%.

TABLE A13 — PERMUTATION TEST

	$\hat{\beta}$	p-value
Sexual harassment	**0.101**	**0.001**
Disability	0.001	0.945
Nationality	0.003	0.800
Age	−0.030	0.011
Race	0.025	0.148

Table A13: This table reports the results of a permutation test in which we re-ran our main specification reassigning treatment to each discrimination type $g(i)$ sequentially in each row (excluding sexual harassment from the control group). The top row, bolded, is our actual treatment. The columns show the coefficient on win, the t-statistic, and the p-value. Religion and Retaliation discrimination bases are excluded because they fail a pre-trends test (i.e., more than two pre-treatment year coefficient estimates are statistically different from zero at the $\alpha = 0.05$ level).

TABLE A14 — $\hat{\beta}^{CE}$ POOLING HOUSING, PUBLIC ACCOMMODATIONS, EDUCATION, AND EMPLOYMENT COMPLAINTS

	All complaints	**Complaints with gender**	
SH × Post	0.130*	0.236***	0.203***
	(0.072)	(0.032)	(0.032)
SH × Post × Female			0.046***
			(0.012)
Unit and Time × State FE	✓	✓	✓
Unit and Time × State × Female FE			✓
N	142,198	84,470	84,384
R^2	0.149	0.091	0.099
Control mean	0.139	0.205	0.207

Table A14: This table is equivalent to Table A10, but we change the sample to include discrimination complaints in employment, housing, public accommodation, and education.

TABLE A15 — $\hat{\beta}^{O}$ POOLING HOUSING, PUBLIC ACCOMMODATIONS, EDUCATION, AND EMPLOYMENT COMPLAINTS

	All complaints	**Complaints with gender**	
SH × Post	0.176***	0.206***	0.207***
	(0.041)	(0.038)	(0.052)
SH × Post × Female			-0.003
			(0.028)
Unit and Time × State FE	✓	✓	✓
Unit and Time × State × Female FE			✓
N	76,491	54,787	54,730
R^2	0.131	0.074	0.081
Control mean	0.139	0.205	0.207

Table A15: This table is equivalent to Table 3, but we change the sample to include discrimination complaints in employment, housing, public accommodation, and education.

D ω calibrations

Although we show in Section 8 that the share of always reporters post #MeToo, ω, does not affect our main qualitative predictions, it may nevertheless be illustrative to consider benchmark values of ω. In this section, we provide two potential such benchmarks.

The share of always reporters can be inferred from changes in the number of cases reported before and after #MeToo, under assumptions about the shares of founded versus unfounded cases before and after #MeToo.[31] We make two sample restrictions to increase consistency. First, we consider only cases in the states contributing to our difference-in-difference coefficients. Second, to avoid the mechanical drop in cases following the Covid-19 pandemic, we consider reporting counts from our data restricting to one year pre-and post-#MeToo. In this subsample, there are 767 sexual harassment cases prior to #MeToo and 957 sexual harassment cases after #MeToo, yielding $957 - 767 = 190$ additional reports in the post period.

Our first calibration method equates the set of induced reporters to precisely be the 190 additional cases. In order for this to hold given our definitions of always and induced reporters, one of the following two limit cases must hold:

- All 767 cases pre-#MeToo were founded, 767 cases post-#MeToo were founded, and 190 cases post-#MeToo were unfounded, or
- All 767 cases pre-#MeToo were unfounded, 767 cases post-#MeToo were unfounded, and 190 cases post-#MeToo were founded.

In either of these cases, the share of always reporters is $\omega_1 := \frac{767}{957} = 0.801$ and the share of induced reporters is $1 - \omega_1 = \frac{190}{957} = 0.199$.

Our second calibration method dispenses with the extreme assumption either that all cases in the pre-period are founded or that all cases in the pre-period are unfounded. Instead, we assume that the share of founded accusations in the pre-period is equal to the share of "won" investigations in the pre-period. This value in our data is 0.194. To account for potential changes caused by #MeToo, we assume that #MeToo lowers the share of founded accusations by 0.014. This value is inspired by Levy and Mattsson (2025), who find that #MeToo decreased sexual *assault* incidence by 1.4 percentage points.[32] Based on these numbers, the share of always reporters is $\omega_2 := \frac{\frac{957(0.194-0.014)}{0.194}}{957} = \frac{0.194-0.014}{0.194} = 0.928$, where $\frac{957(0.194-0.014)}{0.194}$ is the number of always reporters in the post-period, i.e., the maximal number of cases such that precisely 0.194 share are founded. The share of induced reporters is $1 - \omega_2 = 1 - 0.928 = 0.072$.

Our calibrated values are therefore $\omega \in \{0.801, 0.928\}$. We reiterate that these are illustrative, and our qualitative conclusions do not depend on the specific value of ω.

[31] Throughout, we ignore integer constraints implied by data finiteness, and allow for a fractional case to be founded.

[32] Levy and Mattsson's analysis is based on the Campus Climate Survey, which measures the prevalence of sexual misconduct using self-reported data from U.S. university students. Sexual assault and sexual harassment are different, and college students are a selected sample, however, their estimate is possibly the best existing proxy for our setting.

E Data Appendix

E.1 FOIA process

To obtain data on civil complaints of sexual harassment, we filed Freedom of Information Act requests with all US states that have state agencies responsible for handling discrimination complaints and with the federal Equal Employment Opportunity Commission (EEOC). Two states (Arkansas and Mississippi) did not have such a state agency. In its broadest form, our request included information on all closed employment discrimination, public accommodation discrimination, housing discrimination, and education discrimination cases that were filed between June 1, 2010, and June 1, 2023. Requested information would include the respondent's name, filing date, resolution date, alleged basis and issue of discrimination, decision on the complaint (probable cause, no probable cause, etc.), outcome (went to court, dismissed, etc.), and compensation paid to the complainant. Before filing the FOIA request, we would contact someone at the agency whenever possible to ensure our requested data was available and make appropriate changes to our request if needed. Contact information was usually found on the agency's website or directory, specified below.

As of September 2025, we have received usable data containing employment, public accommodation, housing, or education discrimination complaints from 17 states. Our main analysis sample consists of the 10 states that provided employment discrimination data (listed in Table A1). The remaining 7 states provided data on other types of discrimination but not employment discrimination. We use the full 17-state sample in robustness checks in Appendix C.

1. Alaska: Received a spreadsheet after submitting a request via email to hrc@alaska.gov.
2. California: Received a spreadsheet after submitting a request via the online portal.
3. Delaware: Received a spreadsheet after submitting a request via email to dos.foia@delaware.gov.
4. Florida: Received a spreadsheet in PDF format after submitting a request via email to records@fchr.myflorida.com.
5. Hawaii: Received a spreadsheet after submitting a request via email to DLIR.HCRC.INFOR@hawaii.gov.
6. Illinois: Received a spreadsheet in PDF format after submitting a request via email to IDHR.FOIA@illinois.gov.
7. Kentucky: Received a spreadsheet after submitting a request via email to kchr.mail@ky.gov.
8. Massachusetts: Received a spreadsheet after submitting a request via email to cadrao@mass.gov.
9. Michigan: Received a spreadsheet after submitting a request via email to MDCR-FOIA@michigan.gov.
10. Montana: Received a spreadsheet after submitting a request via the online portal.

11. North Carolina: Received a spreadsheet after submitting a request via email to oah.postmaster@oah.nc.gov.
12. North Dakota: Received a spreadsheet after submitting a request via email to labor@nd.gov.
13. New York: Received a spreadsheet after submitting a request via the online portal.
14. South Carolina: Received a spreadsheet after submitting a request via email to information@schac.sc.gov.
15. Texas: Received a spreadsheet in PDF format after submitting a request via the online portal.
16. Washington: Received a spreadsheet after submitting a request via email to records@hum.wa.gov.
17. Wisconsin: Received a spreadsheet after submitting a request via email to OpenRecords@dwd.wisconsin.gov.

We have **received unusable data** from 7 states:

1. Georgia: Received a spreadsheet in PDF format after submitting a request via email to info@gceo.state.ga.us. Data are unusable because there is no data from the pre-period.
2. Maine: Compiled cases from the website after submitting a request via email to foaa@mhrc.maine.gov. Data are unusable because they were provided in unstructured PDF format.
3. Maryland: Received a spreadsheet after submitting a request via email to mccr@maryland.gov. Data are unusable because there is no data from the pre-period.
4. Minnesota: Received a spreadsheet after submitting a request via the online portal. Data are unusable because there is no data from the pre-period.
5. Oklahoma: Received a spreadsheet in PDF format after submitting a request via email to contact@oag.ok.gov. Data are unusable because bases of discrimination were unavailable and were not provided upon further request.
6. Pennsylvania: Received case files after submitting a request via email to RA-RTK_PHRC@state.pa.us. Data are unusable because there is no data from the pre-period.
7. Rhode Island: Received PDFs of case files after submitting a request via email to John.Bogue@richr.ri.gov. Data consisted of 12 observations which were excluded after meeting various sample restrictions.

Our request was **closed/denied by/not responded to within 1 year** the following 24 states:

1. Alabama: Request submitted via the online portal and denied due to lack of state citizenship.
2. Arizona: Request submitted via the online portal.
3. Colorado: Request submitted via email to dora_ccrd@state.co.us and denied on the basis of state law.

4. Connecticut: Request submitted via the online portal.
5. Idaho: Request was closed after the state said that they do not release any files on the basis of confidentiality.
6. Indiana: Request submitted via the online portal and denied due to the range of data requested.
7. Iowa: Request submitted via email to icrc@iowa.gov and denied on the basis of state law.
8. Kansas: Request submitted via email to khrc@ks.gov.
9. Louisiana: Request submitted via email to GovPublicRecords@la.gov and denied on the basis of state law.
10. Missouri: Request submitted via email to mchr@labor.mo.gov and denied on the basis of state law.
11. Nebraska: Request submitted via email to neoc.response@nebraska.gov and denied on the basis of state law.
12. Nevada: Request submitted via email to detrmedia@detr.nv.gov and denied on the basis of state law.
13. New Hampshire: Request submitted via email to humanrights@hrc.nh.gov and denied based on NH Administrative Rule Part 219.04.
14. New Jersey: Request submitted via the online portal and denied on the basis of state law.
15. New Mexico: Request submitted via email to ipra.workforce@dws.nm.gov and closed after the state asked for payment of $6000 to process the request.
16. Ohio: Request submitted via telephone and closed after the state shared that their retention period was only 5 years.
17. Oregon: Request submitted via email to public.records@boli.oregon.gov.
18. South Dakota: Request submitted via the online portal and denied on the basis of state law.
19. Tennessee: Request submitted via email to ask.thrc@tn.gov and denied due to lack of state citizenship.
20. Utah: Request submitted via email to sdanielson@utah.gov and closed after the state said that they did not keep any records of charge filing or resolution dates on specific charges.
21. Vermont: Request submitted via email to human.rights@vermont.gov and denied on the basis of state law.
22. Virginia: Request submitted via the online portal and decided not to pursue after the state asked for a payment of $1097.32 to process the request.
23. West Virginia: Request submitted via contact us page.

24. Wyoming: Request submitted via the online portal and decided not to pursue after the state asked for payment of $440 to process the request.

E.2 Variable definitions

After receiving the complaint data, we categorize each complaint as belonging to one of five jurisdictions: employment, public accommodation, housing, education, and unspecified.

We categorize the basis of discrimination for each complaint into one of seven categories: age, disability, nationality, race, religion, retaliation, and sex. The original basis of discrimination as provided by states is much more specific than these categories. For example, we classify "Sexual harassment" and "Sex: Pregnancy" (North Dakota basis) both as sex.

Each case can have multiple bases. When a complaint is filed on multiple bases, we use the primary basis, usually the first listed. If the primary basis cannot be determined, complaints are classified according to the following order, with later bases taking precedence: age, retaliation, nationality, religion, disability, race, and sex.

In addition to the basis of discrimination, each case is assigned an issue by the state. The issue is the type of action alleged. The most common issues are discharge, harassment, terms/conditions, and wages. A complaint is classified as a sexual harassment case if sexual harassment is listed as an issue and the complaint was filed based on sex in the raw data. Cases that have sexual harassment listed as an issue but were not filed on the basis of sex are excluded from classification.

A complaint is classified as a sex-based case if it was filed on the basis of sex, regardless of the issue or complainant's gender.

Compensation to the complainant is preserved from the raw data and summed together when complainants receive monetary relief at different stages in the complaint process (e.g., at the hearing and in court). Compensation is labeled as missing (instead of 0) when the complainant's case is dismissed or when such information is not available. We created an indicator variable for missing relief to account for such cases.

Complainants are considered to have "won" their case if the raw data indicates that "reasonable cause"/"substantial evidence" of discrimination or the equivalent was found. If the raw data indicates that there was no reasonable cause/substantial evidence of discrimination, complainants are considered to have "lost" their case. In any cases where there is no explicit mention of the agency's/court's findings, including cases that were settled, the winning indicator is marked as missing.

Cases that were settled, dismissed, or went to court are tracked with corresponding indicator variables.

We create an indicator variable for whether a case was resolved after #MeToo.

Some states indicate whether the complainant is female. In some cases, the raw data explicitly

states this, while in others, the raw data provides only the complainant's name and not their sex. In these cases, we use an API called Genderize.io to predict the complainant's most likely gender and then visually inspect the resulting gender. In cases where both types of information are missing from the raw data, this variable is marked as missing.

Outcome Table: This table displays all of the case outcomes in our raw data and how we map each one to a standardized outcome variable. If the outcome was not self-evident from the label, we followed up with state agencies over email to clarify their terminology.

Value	Variable
A1 - Complaint Withdrawn	dismissed == 1
A2 - Complaint not Timely	dismissed == 1
A4 - Complainant Not Available	dismissed == 1
A5 - Failure of CP to Proceed or Cooperate	dismissed == 1
A6 - Complainant to Court	court == 1
A7 - Administrative Dismissal	dismissed == 1
AC: General	dismissed == 1
ADMIN CLSR	dismissed == 1
Adjusted and Withdrawn	settle == 1
Adjusted/Terms of Settlement	settle == 1
Admin Closure	dismissed == 1
Admin Closure Outside of Regular Meeting	dismissed == 1
Admin. Closure	dismissed == 1
Administrative Closure	dismissed == 1
Administrative Closure DWOP	dismissed == 1
Administrative Closure/ DWOP	dismissed == 1
Administrative Closure; DWOP	dismissed == 1
Administrative Dismissal	dismissed == 1
Administratively closed	dismissed == 1
Annulment Issued	court == 1
Annulment Issued (no paper file)	court == 1
B1 - Successful Settlement	settle == 1
B2 - Predetermination Settlement (PDS)	settle == 1
B3 - Not Substantial Evidence (Exception)	win == 0
B4 - Conciliation Finalized	win == 1
C01 - Signed/notarized complaint not returned	dismissed == 1
C1 - Hearing Decision for Complainant	win == 1
C4 - Pre-Hearing Settlement	settle == 1
C5 - Administrative Dismissal	dismissed == 1
CO1 - Post-Investigation Administrative closure	dismissed == 1
CO2 - Post-Investigation Settlement Agreement	settle == 1
CO3 - Post-Investigation Legal Determination of Inability to proceed	dismissed == 1

Value	Variable
COURT-DISMISSED BY JUDGE	court == 1
CP Filed Private Lawsuit	court == 1
CPT withdrawn - Resolved by parties	settle == 1
CPT withdrawn/dropped	dismissed == 1
CPT withdrawn - Cp electing court action	court == 1
Case Settled	settle == 1
Cause	win == 1
Cause Finding Consent Decree	win == 1
Cause Finding Hearing	win == 1
Cause/No Cause	win == 1
Charge Issued	win == 1
Closed - Chapter 478 (removed to court)	court == 1
Closed - Conciliated	win == 1
Closed - Dismissed	dismissed == 1
Closed - Failure to Cooperate	dismissed == 1
Closed - Lack of Probable Cause	win == 0
Closed - Pre-Determination Settlement	settle == 1
Closed - R&A Dismissal	dismissed == 1
Closed - Settled At Hearing	settle == 1
Closed - Unable to Locate Complainant	dismissed == 1
Closed - Withdrawn	dismissed == 1
Closed - Withdrawn With Settlement	settle == 1
Closed Because Trial Has Begun	court == 1
Complainant Elected Court Action	court == 1
Complainant Failed To Cooperate	dismissed == 1
Complainant Filed Suit	court == 1
Complaint Dismissed	dismissed == 1
Complaint Rejected	dismissed == 1
Complaint Withdrawn	dismissed == 1
Complaint Withdrawn by Complainant After Resolution	settle == 1
Complaint Withdrawn by Complainant Without Resolution	dismissed == 1

Value	Variable
Conc	win == 1
Conc by ENFORCEMENT	win == 1
Conciliated	win == 1
Conciliated Settlement	win == 1
Conciliated/Dismissal w/Pred.	win == 1
Conciliation	win == 1
Conciliation Closing Issued	win == 1
Conciliation Closing Issued (no paper file)	win == 1
Conciliation Failure	win == 1
Conciliation/settlement successful	settle == 1
Cp failed to co-operate	dismissed == 1
Cp not available	dismissed == 1
D3 - Appeal - Affirmed	win == 0
Determination Dismissing Complaint Issued	dismissed == 1
Determination by Another Agency	court == 0; dismissed == 0
DF	win == 1
DWOP	dismissed == 1
Decline to File: Exercise of Discretion	court == 0; dismissed == 0
Dismissal	dismissed == 1
Dismissal without Pred.	dismissed == 1
Dismissal w/out prejudice	dismissed == 1
Dismissal w/Pred.	dismissed == 1
Dismissal w/Settlement	dismissed == 1; settle == 1
Dismissal without Pred./NPC	win == 0
Dismissal without Prejudice/Admin Closure	dismissed == 1
Dismissed for Lack of Jurisdiction	dismissed == 1
E03 - Untimely for MDCR and / or EEOC	dismissed == 1
EEOC=NJ & NDHRA=NPC	win == 0
Expired	dismissed == 1
FHAP judicial consent order	settle == 1
FHAP judicial dismissal	dismissed == 1
Failure to Cooperate	dismissed == 1
Failure to Locate	dismissed == 1
Failure to Proceed	dismissed == 1

Value	Variable
I01 - Insufficient evidence - adjusted	win == 0
I02 - Insufficient evidence	win == 0
I03 - Settlement Agreement	settle == 1
I04 - Unable to locate claimant	dismissed == 1
I05 - Claimant failure to cooperate	dismissed == 1
I06 - Wrong respondent	dismissed == 1
I07 - Claimant in court	court == 1
I11 - Withdrawn adjusted	settle == 1
I12 - Withdrawn - pursue in Court - no RTS	court == 1
I13 - Withdrawn - pursue in Court - with RTS	court == 1
I14 - Withdrawn - Not interested in pursuing	dismissed == 1
I16 - Decided by Court - no ruling on merits	court == 1
I17 - Decided by Court- w/adjustment	win == 1
I18 - Decided by Court – no adjustment	court == 1
INVESTIGATED BY HUD	win == 0
INVESTIGATED- NO COURT	court == 0
Insufficient Evidence	win == 0
Intake Closure	dismissed == 1
Investigated and Dismissed	dismissed == 1
Investigated and Dismissed - Accusation Withdrawn	dismissed == 1
Investigated and Dismissed - Insufficient Evidence	dismissed == 1
Investigated and Dismissed - Insufficient Evidence and NJI	dismissed == 1
Investigated and Dismissed - Insufficient Evidence and NJI - CP Failed to Cooperate	dismissed == 1
Investigated and Dismissed - Insufficient Evidence and NJI - CP Failed to Respond to 30-Day Letter	dismissed == 1
Investigated and Dismissed - Insufficient Evidence and NJI - CP Filed Suit	court == 1
Investigated and Dismissed - Insufficient Evidence and NJI - Closed Due To Court Decision	court == 1
Investigated and Dismissed - Insufficient Evidence and NJI - Respondent unavailable	dismissed == 1
Investigated and Dismissed - No Basis to Proceed	dismissed == 1
Investigated and Dismissed - Withdrawn - Intend to file a lawsuit	court == 1
Investigated and Dismissed - Withdrawn - Other reasons	dismissed == 1

Value	Variable
Investigated and Dismissed - Withdrawn - Rescind	dismissed == 1
Investigated and Dismissed - Withdrawn - Resolved between parties	dismissed == 1
Investigation not Completed within Statutory Time Limit	dismissed == 1
Judgment - Unfavorable	win == 0
L01 - Post-Investigation Administrative closure	dismissed == 1
L02 - Post-Investigation Settlement Agreement	settle == 1
L03 - Post-Investigation Legal Determination of Inability to proceed	dismissed == 1
Lack of Substantial Evidence	win == 0
M01 - Settlement Agreement	settle == 1
M02 - Withdrawn adjusted	settle == 1
M03 - Withdrawn - not interested in pursuing	dismissed == 1
M1 - Mediation Successful	settle == 1
M2 - CP Withdrawn with Settlement	settle == 1
M3 - Complaint Withdrawn in Mediation	settle == 1
M4 - Mediatin Predetermination Settlement	settle == 1
MEDIATED BY COUNSEL	win == 1
Mediation/Settlement	settle == 1
NPC	win == 0
NPC Reactivated by HUD over 120 days old	win == 0
NPC W/OUT PREJUDICE	win == 0
NPC ; NPC on Recon	win == 0
NPC ; Recon DENIED	win == 0
NPC ; See Notes	win == 0
NPC Dismissal with Prejudice	win == 0
NPC w/out prejudice	win == 0
NPC/ Withdrawal	win == 0
NPC/ Withdrawal WITHOUT Settlement	win == 0
NPC/Order to Vacate & Set Aside & to Grant Withdrawal/Withdrawal	win == 0
NPC/Recon Pending	win == 0
NPC/Reconsideration	win == 0
Negotiated Settlement	settle == 1
No Basis to Proceed (For Filing Purposes Only)	dismissed == 1

Value	Variable
No Cause	win == 0
No Cause Determination	win == 0
No Cause Finding	win == 0
NO CAUSE - insufficient evidence	win == 0
No Follow-Up Info Rcvd from Requestor	dismissed == 1
No Probable Cause	win == 0
No Probable Cause Determination Issued	win == 0
No Reasonable Cause	win == 0
No cause determination	win == 0
Notice of Right to Sue	court == 1
Notice of Rights	win == 0
PC Determination Issued / Post Cause Conciliation	win == 1
PC Determination Issued / Conciliated	win == 1
PC Determination Issued / Conciliation	win == 1
PC Determination Issued / Dismissed by Circuit Court	win == 1
PC Determination Issued / Withdrawal WITH Settlement	win == 1
Post Cause Conciliation	win == 1
Pre-finding Settlement	settle == 1
Probable Cause	win == 1
Probable Cause Determination Issued	win == 1
Probable Cause Found	win == 1
Processing Waived to Another Agency	court == 0; dismissed == 0
RC - Resolution Conference Closure	settle == 1
Reasonable Cause Recommendation - Settled	win == 1
Reasonable Cause Recommendation - Withdrawn	win == 1
Reasonable remedy refused	court == 0; dismissed == 0
Relief covered by Ct order/consent decree	settle == 1
Resolved between Parties	settle == 1
Right to Sue	court == 1
Right to Sue Issued	court == 1
Right to Sue—Intake Closure	dismissed == 1

Value	Variable
Serve Annulment (Post-Investigation)	court == 1
Serve Final Order Dismissing Complaint	dismissed == 1
Serve Order After Hearing: Dismissing Complai	win == 0
Serve Order After Hearing: Sustaining Complain	win == 1
Serve Order After Stipulation of Settlement	settle == 1
Settled after Charge Issued	win == 1
Settled by Appeals	settle == 1
Settled by DRD and Legal: Mandatory Mediation	settle == 1
Settled by DRD and Legal: Post-civil Complaint	settle == 1
Settled by DRD: Post Appeal	settle == 1
Settled by DRD: Voluntary Mediation	settle == 1
Settled by Enforcement	settle == 1
Settled by Legal: Post- civil Complaint	settle == 1
Settled by Legal: Pre- Filing	settle == 1
Settled by Mediation: Voluntary	settle == 1
Settled with Respondent (specific complaint not sustained; other inequities remedied)	settle == 1
Settlement - Settled by Enforcement	settle == 1
Settlement - Settled by Mediation - Mandatory 1038	settle == 1
Settlement - Settled by Mediation - Voluntary	settle == 1
Settlement - Settled with Respondent - Pre- Civil	settle == 1
Split Decision: PC/LOPC or PC/LOJ	win == 1
Stay Substantial Evidence	win == 1
Successful Conciliation	win == 1
Transfer to EEOC (Closed at Commission)	dismissed == 1
Transferred to Circuit Court	court == 1
Unable to Locate Respondent	dismissed == 1
Unable to locate complainant	dismissed == 1
Unable to Serve	dismissed == 1
Untimely filed	dismissed == 1
VRA	settle == 1
WD	dismissed == 1
WDB	settle == 1

Value	Variable
Withdrawal	dismissed == 1
Withdrawal with Benefits	settle == 1
Withdrawal With Settlement	settle == 1
Withdrawal Without Settlement	dismissed == 1
Withdrawal w/o Predjudice	dismissed == 1
Withdrawal with Resolution/Settlement	settle == 1
Withdrawal without Resolution/Settlement	dismissed == 1
Withdrawn	dismissed == 1
Withdrawn With Resolution	settle == 1
Withdrawn Without Resolution	dismissed == 1
Withdrawn Request â Housing Only	dismissed == 1
Withdrawn Request â Housing Only - Intend to file a lawsuit	court == 1

Note: We do not include duplicate outcomes, including those differing only by spelling errors or capitalization. Double-mapped outcomes (e.g., court == 0; dismissed == 0) represent cases where administrative records allow us to rule out specific outcomes but are insufficient to classify the case further.