Sexual Misconduct and Scientific Production

Manuela R. Collis *
University of Toronto

This version: March 26, 2026 — Please click here for latest version

Abstract

Although sexual misconduct in the workplace has complex and lasting consequences for directly affected individuals, its broader organizational implications remain less well understood. Using a newly assembled dataset of 1,294 publicly documented sexual misconduct cases across U.S. universities—recording both the year in which sexual misconduct occurred and the year it became public—I examine how these incidents affect departmental scientific productivity. I employ coarsened exact matching and a staggered difference-in-differences design to compare departments that experienced subsequently publicized misconduct incidents with matched control departments. I find that sexual misconduct does not show a discernible effect on departmental productivity when it occurs, but public reporting reduces publications by 0.1 per faculty member annually, equivalent to nine fewer publications over five years for a median department of 18 members. These findings indicate that organizational costs arise specifically from public disclosure rather than from misconduct itself, suggesting that protecting victims while maintaining productivity may require differentiated policy approaches as institutions navigate competing legal frameworks, ethical obligations, and performance-related concerns. This distinction helps explain both why social pressures can transform misconduct from a human resource concern into a strategic organizational challenge and why firms may prioritize confidentiality strategies to the potential detriment of victims.

*Collis: Rotman School of Management, 105 St. George Street, Toronto, ON M5S 3E6, Canada (email: manuela.collis@rotman.utoronto.ca). I am deeply grateful to my committee, András Tilcsik, Avi Goldfarb, Nicola Lacetera, and Sarah Kaplan, for their invaluable support and feedback throughout this project. I thank Megan McGarvie and Laura Doering for helpful comments. I thank Afyaa Enaya, Yiwen (Catherine) Ding, Rahdin Salehian, Yashasvini Singh, and Ruth Zhang for excellent research assistance. This paper benefited from discussions at the CCC Conference, the AOM STR Dissertation Consortium, the Rotman Strategy department, the Field Lab group, and seminar participants at UC Berkeley, Bocconi University, Carnegie Mellon University, IE Business School, NYU Stern. I gratefully acknowledge funding support from TD-MDAL, Rotman. All errors are my own. The study was preregistered on May 10, 2025, and the preregistration can be accessed under https://osf.io/zx3gu.

1 Introduction

Sexual misconduct in the workplace, which encompasses a spectrum of harmful behaviors including harassment, assault, and jokes or comments that devalue individuals on the basis of their sex (Equal Employment Opportunity Commission, 1980), is widespread and has complex and lasting consequences for affected individuals. Nearly 60% of women and 27% of men report having personally experienced unwanted sexual advances or sexual harassment, either verbal or physical, in workplace or non-workplace settings (Graf, 2018). Experiencing sexual misconduct can have severe personal consequences: it diminishes job satisfaction, heightens the desire to leave one's position, and undermines psychological well-being, potentially leading to anxiety, stress disorders, and depression (Schneider, Swan and Fitzgerald, 1997; Cortina and Berdahl, 2008; Fitzgerald and Cortina, 2018). Several studies show that it drives women away from organizations (Folke and Rickne, 2022; Adams-Prassl et al., 2024; Collis and Van Effenterre, 2024; Batut, Coly and Schneider-Strawczynski, 2022). Although researchers have made important strides in documenting the consequences of sexual misconduct on individual workers, much less is known about how it affects organizational outcomes such as productivity.

Paradoxically, organizations invest substantial resources in preventing sexual misconduct while making efforts to protect perpetrators when incidents occur. Many companies try to follow best practices by, for example, mandating anti-sexual-harassment training (Antecol and Cobb-Clark, 2003; Kalev, Dobbin and Kelly, 2006; Dobbin and Kelly, 2007; Dobbin and Kalev, 2019). In fact, states such as New York, Massachusetts, and California require such programs by law. However, when misconduct occurs, organizations routinely implement nondisclosure agreements and retain accused employees (Adams-Prassl et al., 2024; Barmes, 2023). A possible explanation is that leaders are more concerned about the organizational consequences of public exposure than about the moral and legal risks of sexual misconduct itself. To explore this idea, this study examines the productivity consequences of sexual misconduct in organizations both at the time the misconduct occurs and at the time an incident is publicly reported.

Studying sexual misconduct in organizations is challenging. Sexual misconduct is systematically underreported, and reported incidents are often not recorded (Fitzgerald, Swan and Fischer, 1995; Boudreau et al., 2023; Dahl and Knepper, 2021; Cheng and Hsiaw, 2022), making data hard to obtain. Moreover, granular personnel and productivity data are required to study its effects on overall productivity.

To overcome these challenges, I turn to university departments. This setting has several key advantages. First, although imperfect, data on university sexual misconduct cases exist. The greatest challenges in studying sexual misconduct are that it is not systematically measured, and the data that are available, for example, from the U.S. Equal Employment Opportunity Commission (EEOC), are de-identified, preventing us from studying longitudinal outcomes of organizations and their members. U.S. Universities are governed under Title IX and are thus systematically measured (relatively speaking). Second, the data are more easily accessible, for example, through freedom of information requests. Third, one of a university department's key outputs—scientific publications—is an established measure and provides rich information about productivity. Moreover, university departments are central sites of scientific production and intellectual innovation, with far-reaching implications for knowledge creation and broader economic outcomes (Agrawal and Henderson, 2002; Azoulay, Graff Zivin and Wang, 2010; Stephan, 2010; Chandra and Xu, 2025).

In this empirical setting, I examine *how sexual misconduct incidents in university departments affect scientific production* by building a panel dataset for over 5,000 university departments across nine disciplines between 1980 and 2024. To construct the treated departments, I use a comprehensive dataset of all sexual misconduct cases at universities in the United States that were publicized between 1980 and 2024. To study the relationship between sexual misconduct and productivity, I restrict my sample to research-focused universities where the accused individual was a faculty member, resulting in 359 treated university departments. To construct a set of comparable control departments, I use a list of unique research-focused universities based on data from the Integrated Postsecondary Education Data System (IPEDS) and the Carnegie Classification. I then randomly select nine departments per university as

the controls. This results in a total of 5,076 control university departments.

For the 359 sexual misconduct incidents, I hand-collect data from over 1,300 lawsuits, news articles, or proprietary documents shared through online document libraries, which allow me to establish two crucial points of time: the year when sexual misconduct started to occur and the year when public reporting began. I use these two points of time as two different "treatment" times to establish the causal impact on productivity of (1) the occurrence of sexual misconduct and (2) the public reporting of sexual misconduct.

To construct a suitable comparison group for departments where sexual misconduct was publicly reported, I pair coarsened exact matching (CEM) with a staggered difference-in-differences (DiD) event study design. University departments are matched on five covariates: the presence of state laws on discrimination and harassment, whether the university is public or private, the academic discipline, the research intensity of the department, and the department's size.

The results indicate that the *public reporting* of sexual misconduct incidents decreases university departments' overall productivity, but that the *occurrence* of sexual misconduct does not. Specifically, the public reporting of sexual misconduct reduces the average number of publications per department member per year by 0.1 ($p < 0.05$) within the first five years after the incident becomes public. In other words, a median department of 18 faculty members will have produced nine fewer publications five years after the incident becomes public than its comparison departments. However, the average number of publications per department member where sexual misconduct occurred is indistinguishable from that in control departments. This null effect at the average masks some heterogeneity, however. I find that in departments where withdrawal from research is possible and costless—for example, in clinical medicine, by switching to patient-focused activities without financial penalty—sorting does occur, and department members shift away from research activities.

My main findings, which indicate a substantial distinction between contained and spillover effects, are consistent with stigma-based and cognitive mechanisms. When sexual misconduct remains private, stigma, fear of retaliation, and privacy concerns typically contain its orga-

nizational impact. Specifically, my finding that the occurrence of sexual misconduct has no discernible effect on overall productivity suggests that directly affected individuals may avoid discussing the stigmatized event with others because of self-blame, shame, and fear of negative judgment (Rowe, 1996; Cortina and Magley, 2003; Rehg et al., 2008; Dobbin and Kalev, 2019; Dahl and Knepper, 2021). Thus, knowledge about misconduct incidents appears to remain limited to directly affected individuals and those with close ties to them.

However, when sexual misconduct becomes public, different mechanisms emerge. Public reporting creates reputational damage and productivity losses that extend beyond the direct victims, consistent with stigma by association (Lange, Lee and Dai, 2011), in which publicized misconduct negatively affects individuals with observable ties to the accused through department membership (McDonnell and Werner, 2016; Barnett and King, 2017; Janney and Gove, 2017) or shared social group characteristics (Greve, Palmer and Pozner, 2010; Jensen, 2006).

These findings suggest that organizational costs arise from public disclosure rather than from misconduct itself. Thus, organizations may face a perverse incentive to keep misconduct hidden rather than working to prevent it and protect victims. This helps explain why organizations often appear reluctant to provide full transparency and fail to achieve justice for victims: they must balance moral imperatives against significant organizational consequences. As social pressure increasingly demands transparency, sexual misconduct is transforming from a human resource concern into a strategic organizational challenge.

2 Sexual Misconduct in Organizations

Sexual misconduct refers to unwelcome sexualized conduct and consists of unwanted sexual attention and sexual coercion (Cortina and Areguin, 2021; Cheng and Hsiaw, 2022). Unwanted sexual attention includes unwelcome and sometimes traumatizing expressions of sexual interest. Examples include unwanted sexualized language, nonconsensual touching, forcible kissing, and sexual assault (Fitzgerald et al., 1988; Fitzgerald, 2019). Sexual coercion, on the other hand, is the act of making employment (i.e., getting and keeping a job) or professional opportu-

nities (e.g., promotions, grants, social recognition) contingent on sexual cooperation (Cortina and Areguin, 2021). Sexual misconduct often goes hand in hand with gender harassment, which encompasses sexual jokes and acts that devalue individuals on the basis of their sex (Leskinen, Cortina and Kabat, 2011; Leskinen and Cortina, 2014). This definition is more extensive than the legal definition often found in the literature (Folke and Rickne, 2022; Collis and Van Effenterre, 2024), as legal frameworks depend on local legislative frameworks and focus on the harms caused to the victim, whereas in this study, I examine the effects of sexual misconduct on organizations (Cheng and Hsiaw, 2022).

Sexual misconduct is prevalent across a variety of organizations. In a nationally representative random-digit-dial telephone survey in the United States, 47% of respondents stated that they had experienced sexual harassment at work (Rospenda, Richman and Shannon, 2009). The organizational impact of these incidents is amplified by the fact that they are typically not isolated events. In one survey, 66% of respondents who had recently experienced such misconduct reported that it happened to them more than once (Morral, Gore and Schell, 2015).

Sexual misconduct produces significant adverse outcomes for victims, both psychological and occupational. It is positively correlated with depression, anxiety, stress, and post-traumatic stress disorder (Schneider, Swan and Fitzgerald, 1997; Cortina and Berdahl, 2008; McLaughlin, Uggen and Blackstone, 2017). Evidence suggests that it leads to decreased job satisfaction, an increased chance of withdrawal, and an increased likelihood of leaving (Schneider, Swan and Fitzgerald, 1997; Sims, Drasgow and Fitzgerald, 2005; Cortina and Berdahl, 2008; McLaughlin, Uggen and Blackstone, 2017; Adams-Prassl et al., 2024).

These sexual misconduct incidents can have profound consequences for directly affected individuals. Victims can experience psychological effects including increased stress and depression, behavioral consequences such as absenteeism, and lasting damage to career progression (Schneider, Swan and Fitzgerald, 1997; Cortina and Berdahl, 2008; Fitzgerald and Cortina, 2018). These effects are severe enough to drive women from organizations entirely (Folke and Rickne, 2022; Adams-Prassl et al., 2024; Batut, Coly and Schneider-Strawczynski, 2022; Collis

and Van Effenterre, 2024).

Although many factors contribute to the occurrence of sexual misconduct in the workplace, such as individual characteristics (Pryor, 1987; Pryor, LaVite and Stoller, 1993) or status (Berdahl, 2007), a large body of work shows that organizational climate, gender imbalance, and power imbalance play important roles. For instance, organizations that maintain a climate of respect and fairness that does not tolerate offensive behavior have fewer reported cases of sexual misconduct (Fitzgerald, Hulin and Drasgow, 1994; Pryor, LaVite and Stoller, 1993; Willness, Steel and Lee, 2007). Moreover, being in the gender minority within an occupation is strongly correlated with experiencing sexual misconduct (Mansfield et al., 1991; Welsh, 1999; Folke and Rickne, 2025; Subramani and Gorbatai, 2025). Both men and women have heightened risks of sexual harassment from colleagues and supervisors in occupations numerically dominated by the opposite gender (Folke and Rickne, 2025). Similarly, power asymmetries, whether due to differences in seniority or economic vulnerability, contribute to the risk of experiencing sexual misconduct (Fitzgerald, Hulin and Drasgow, 1994; Ilies et al., 2003; Popovich and Warren, 2010; Hershcovis et al., 2017; Adams-Prassl et al., 2024). Moreover, power itself can trigger misconduct against women who achieve it, as harassment is often used as a sanction for violating prescriptive gender norms at work (McLaughlin, Uggen and Blackstone, 2012).

Sexual Misconduct Incidents and Organizational Performance

Although research documents the consequences of sexual misconduct for those directly targeted, there is reason to expect them to extend further. Misconduct is rarely experienced in isolation; it unfolds in organizational contexts where the impact of its occurrence and handling may reverberate beyond the immediate victim (e.g., Acquadro Maran, Varetto and Civilotti, 2022; Trawalter et al., 2022). Moreover, if misconduct is revealed publicly, it may have additional implications for members of the organization (Cheng et al., 2024). By shifting attention to these broader potential dynamics, we can begin to capture the organizational consequences of sexual misconduct more comprehensively. This focus also resonates with the broader lit-

erature on organizational wrongdoing and deviance, which demonstrates that individual acts of misconduct can have significant implications for organizational performance that extend far beyond the directly affected parties and the immediate perpetrator (see Alexander, 1999; Karpoff, Lee and Vendrzyk, 1999; Karpoff, Lee and Martin, 2008).

As I discuss in this section, some organizational consequences of sexual misconduct may arise directly from its occurrence. These include increased employee sorting and internal disruptions associated with investigations, remedial actions, and related processes. Beyond these immediate effects, public reporting of misconduct, which is often unexpected and abrupt, might generate additional ripple effects. Public reporting may compel the organization to focus on reputation management, and its members will need to grapple with implications of both the misconduct and the resulting public discourse. In addition, when an incident of misconduct becomes public knowledge, external audiences such as suppliers, investors, and other exchange partners may distance themselves from the focal organization, with potential adverse effects on its performance.

Organizational Effects of Occurrence

There are reasons to believe that the occurrence of sexual misconduct might affect productivity through employee sorting and organizational measures. Evidence suggests that victims share their experiences with coworkers of the same gender and warn them to keep them safe through "whisper networks" (Meza, 2017; Johnson, 2023). Learning about a coworker's experience with sexual misconduct at the workplace is shown to affect task conflict and performance (Raver and Gelfand, 2005). Additionally, although evidence is sparse, some research suggests that witnessing sexual misconduct is associated with negative emotional and psychological consequences, and it may even trigger physical harassment toward the bystander (Hitlan, Schneider and Walsh, 2006; Dionisi and Barling, 2018; Flecha, 2021; Acquadro Maran, Varetto and Civilotti, 2022). Given this body of evidence, it is possible that the occurrence of sexual misconduct affects organizational performance via collaboration decisions and worker disruption.

Moreover, reports of sexual misconduct may trigger organizational investigations and audits, which may lead to remedial measures. These investigations often involve interviewing the parties and witnesses involved (Bedera, 2024) and may culminate in organization-wide interventions, such as extensive sexual misconduct training programs, new grievance policies and procedures, audits, and other bureaucratic measures (see, e.g., Cortina and Berdahl, 2008; Dobbin and Kelly, 2007). Taken together, processing and addressing sexual misconduct complaints may be disruptive and time-intensive and can divert attention and resources from organization members' core functions.

However, these effects may be constrained by barriers to the intra-organizational flow of information. Many sexual misconduct incidents remain unreported (Boudreau and Kaushik, 2023; Sockin and Sojourner, 2023), as affected individuals often prefer to keep information about the incidents to themselves (Rowe, 1996). Victims and witnesses may be particularly reluctant to share information or report incidents because of a fear of stigma and retaliation (Cortina and Magley, 2003; Rehg et al., 2008; Dobbin and Kalev, 2019; Dahl and Knepper, 2021). Even if an incident is reported, organizational procedures are typically designed to maintain confidentiality and minimize the spread of information. For instance, the EEOC states that "employers will protect the confidentiality of harassment complaints to the extent possible" (Feldblum and Lipnic, 2016, p. 38). In addition, many organizations reinforce secrecy through formal confidentiality policies and nondisclosure agreements that prohibit parties from discussing incidents and investigations once concluded.

Even when some organization members eventually learn of a misconduct case, for instance, through whisper networks (Johnson, 2023; Meza, 2017), pluralistic ignorance may keep the information contained (Allport, 1924). Several organization members may know about a transgression yet remain uncertain whether others are aware. Believing that they might be the only ones who know, individuals may hesitate to act, creating a barrier to coordination. In addition, the threat of a potentially damaging scandal may discourage organization members from discussing incidents openly (see Adut, 2005). As a result, information about sexual misconduct incidents may not travel far within an organization, and thus their organizational

consequences may be limited.

Organizational Effects of Public Reporting

Sexual misconduct incidents may become public knowledge through news coverage or whistleblower disclosures. Such public revelations may create distinct dynamics both inside and outside the organization. Internally, public reporting can lower coordination barriers to collective action. After an incident is publicly reported, it becomes common knowledge within the organization. Given the barriers to information flow discussed earlier, many organization members may learn of the case for the first time. Thus, public revelations eliminate ignorance of the incident and resolve the pluralistic ignorance surrounding it. In addition, public revelation can lead to disruptive publicity (Adut, 2005), making inaction within the organization difficult to sustain.

Regardless of prior knowledge, the public disclosure of sexual misconduct incidents is likely to trigger sensemaking among organization members as they seek to interpret and process the event. Sensemaking, defined as an effort to understand issues or events that are novel, ambiguous, or contrary to expectations (Maitlis and Christianson, 2014), is a common response to crisis situations (Wiesenfeld, Wurthmann and Hambrick, 2008) and arises when unexpected events disrupt routine activities. Through this process, individuals construct plausible explanations by selecting and interpreting cues through existing mental frameworks (Weick, 1988, 1995; Weick, Sutcliffe and Obstfeld, 2005; Maitlis and Sonenshein, 2010). Because sensemaking is an inherently social process (Allport, 1924; Maitlis and Christianson, 2014), public disclosure of sexual misconduct may produce complex and disruptive cognitive and social processes across the organization. Such disruptions may be especially consequential in knowledge-intensive settings, where organizational outcomes depend directly on cognitive performance (Blackler, 1995).

Furthermore, public revelation stigmatizes the organization itself (Adut, 2005; Hudson, 2008), as seen in other domains such as bankruptcy (Sutton and Callahan, 1987; Neu and Wright, 1992; McKinley, Ponemon and Schick, 1996) and doping scandals in sports (Yenkey

and Palmer, 2025). Organization members also become tainted by association (McDonnell and Werner, 2016; Barnett and King, 2017; Janney and Gove, 2017; Zhang et al., 2021). Stigmatization, particularly when morally grounded, often provokes strong emotions and social punishment, although the extent of the resulting disrepute and sanctioning depends on perceived blameworthiness (Jones, 1984; Gomulya and Boeker, 2016; Bruyaka, Philippe and Castañer, 2018).

Organizations and their members may respond to stigmatizing events through strategies such as covering (controlling the narrative), withdrawing from relationships, or passing (maintaining normalcy; Goffman, 1963; Page, 1984; Sutton and Callahan, 1987). In the case of publicized sexual misconduct, however, passing and withdrawal are often untenable, leaving narrative control as the primary option. This typically requires extensive communication with collaborators and partners. While individual members engage in sensemaking and narrative management at the personal level, organizations confront the imperative of managing reputational crisis. This task is fraught: traditional strategies for addressing scandals can prove inadequate or even counterproductive (Page, 1984; Sutton and Callahan, 1987; Cheng et al., 2024). Concealment or deflection may appear as complicity, reframing is infeasible, and denying responsibility risks being perceived as victim blaming. As a result, transparency and the acceptance of responsibility are often the most viable strategies, although these approaches are time-consuming and resource intensive, potentially diverting attention and resources away from core organizational functions.

Externally, audiences that serve as key gatekeepers to resources and organizational success often respond to publicized sexual misconduct by distancing themselves from the stigmatized organization. Because stigma spreads through visible ties, suppliers, investors, collaborators, and customers may withdraw or downgrade their support (Goffman, 1963; Pontikes, Negro and Rao, 2010; McDonnell and Werner, 2016; Barnett and King, 2017; Janney and Gove, 2017). Such distancing is documented across diverse contexts: firms entering Chapter 11 bankruptcy experience withdrawal from suppliers and degraded product quality Sutton and Callahan (1987), major adverse events reduce collaborations (McDonnell, Odziemkowska and

Pontikes, 2021), public reporting in stigmatized industries increases divestiture (Durand and Vergne, 2015), and stigmatizing revelations diminish customer engagement (Piazza and Jourdan, 2018). Stigma by association can also operate at highly granular levels: blacklisted workers depress their colleagues' employment prospects (Pontikes, Negro and Rao, 2010), and stigmatized events can even curtail "ceremonial" citations in academia (Azoulay, Bonatti and Krieger, 2017). Taken together, these findings suggest that external distancing may be one way through which publicized misconduct can affect organizational outcomes, likely unfolding simultaneously with the internal processes described above.

3 Empirical Design, Data, and Estimation

Studying sexual misconduct comes with several challenges. It is systematically underreported, and reports often go unrecorded (Fitzgerald, Swan and Fischer, 1995; Boudreau et al., 2023; Dahl and Knepper, 2021; Cheng and Hsiaw, 2022). Moreover, granular data on both personnel and productivity are required to study organizational outcomes and internal dynamics. To overcome these challenges and test the idea that stigma associated with sexual harassment affects organizational behavior, I turn to university departments. This setting has several key advantages. First, although imperfect, data about sexual misconduct cases exist. The challenge in studying sexual misconduct is that it is not systematically measured, and the data that do exist, for example, from the EEOC, are de-identified, making it difficult to study longitudinal outcomes of organizations and their members. Universities are governed under Title IX, so their outcomes are measured systematically (relatively speaking). Crucially, their data can be accessed easily, for example, through freedom of information requests. Second, one of a university department's key outputs—the rate of scientific publications—are an established measure that yields rich information about productivity and collaboration patterns. Moreover, university departments are central sites of scientific production and intellectual innovation, with far-reaching implications for knowledge creation and broader economic outcomes (Agrawal and Henderson, 2002; Azoulay, Graff Zivin and Wang, 2010; Stephan, 2010;

Chandra and Xu, 2025).

To study the impact of sexual misconduct incidents on organizational and societal outcomes in academia, I construct a dataset of productivity outcomes.[1] The departments are listed in Appendix A. Treated departments are defined as those that have experienced a sexual misconduct case that became public.

3.1 Treated Research Departments

To construct the dataset of treated research departments, I use the Academic Sexual Misconduct Database (ASMD; Libarkin, 2024) to define a list of academics who have been associated with an exposed sexual misconduct incident. The database, as accessed on November 19, 2024, contains 1,294 incidents.[2] These are all sexual misconduct incidents in academia that became public through a lawsuit or media reporting and include resolved and ongoing incidents. Throughout this paper, a "publicized" incident refers to one that received newspaper coverage reporting either the start of an investigation or lawsuit, subsequent updates, or its conclusion. The incidents involve faculty, administrators, coaches, and other staff. These cases were reported between 1980 and 2024. For the purposes of this study, I focus on *resolved incidents* at *research-focused* institutions committed by *faculty*. This removes cases that may have ongoing investigations and for which only the general university is known. This also removes technical colleges and U.S. military academies. It also excludes misconduct committed by a coach or administrator, such as a dean, an athletic director, or an executive director, and cases in which the perpetrator's role has not been made public.[3] This reduces the sample to

[1] I define departments by following Clarivate's Core Collection, which contains nine disciplines and 256 categories (https://mjl.clarivate.com/help-center), and supplement these with SCImago categories when necessary (https://www.scimagojr.com/journalrank.php). For the purposes of this work, *departments* are defined as scholars who publish in a list of journals that corresponds to their department. Clarivate and SCImago call them categories.

[2] The raw dataset contains two duplicates, which I removed.

[3] A university is considered research focused if it has ever in its history been classified by the Carnegie Classification in one of the following three categories: "Research 1 - Very High Research Spending and Doctorate Production," "Research 2 - High Research Spending and Doctorate Production," and "Research Colleges and University." I conduct a manual search and classification for the 57 universities not included in the Carnegie Classification. This removes 575 universities, which are mostly technical colleges or U.S. military academies. Among the research-active universities, I remove ongoing cases (173) and cases for which both the individual and the discipline are unknown (83). Lastly, I remove all incidents committed by a coach, administrator, or

359 incidents.

To construct the outcome variables, I collect information about the institution and department at which the respective incidents occurred: publications, citation information, reference information, and abstracts for each institution–department pair. I collect publication data from OpenAlex, an open-source database that provides access to over 200 million publication records by combining databases such as Microsoft Academic Graph, Crossref, ORCID, Unpaywall, Pubmed, Pubmed Central, and ISSN (Priem, Piwowar and Orr, 2022).[4] OpenAlex claims to have superior linkage of author, institution, and publication information compared with alternative data providers such as Web of Science, Lens.org, and Dimensions.

Each of the 359 incidents in the dataset is manually examined and assigned to a department using four pieces of information: Libarkin's assignment of academic discipline, online news articles, an algorithm based on the faculty's publications (in cases where they could be linked to OpenAlex), and the department affiliation reported on their publications (in cases where they could be linked to OpenAlex).[5]

The departments are constructed through topic categories defined by Clarivate and SCImago Journal Rankings, each of which is linked to a corresponding journal list.[6]. Some examples of topic categories are "Political Science," "Geology," and "Marine Biology." There are a total of 527 distinct topic categories, which are listed in Appendix A. My dataset consists of 142 unique departments. Each of these is linked to a journal list, which is used to track publications for a given department.

The 359 incidents took place at 139 unique institutions (one institution is unknown). All 139 institutions could be matched to OpenAlex.[7] Of the 359 incidents, I can link 347 accused

unknown (126).

[4]More information can be found here: https://openalex.org/about

[5]For academics in the life sciences, these assignments were carried out by research assistants with experience in the life sciences; I assigned the individuals in the social sciences to departments myself.

[6]The journal list corresponding to the SCImago topic category "Economics and Econometrics," for example, consists of 166 titles including *American Economic Review*, *Economics of Energy and Environmental Policy*, and the *NBER Macroeconomics Annual* book series.

[7]I employ automatic matching by institution name. Each institution that did not return a unique match is addressed by performing a manual search or removing false positives, as necessary. I also remove universities outside the United States. For example, Duke University (https://openalex.org/I170897317) has a joint university with Wuhan University called Duke Kunshan University, which is located in China (https://openalex.org/I4210159968). I remove that university because it is not located in the United States

and publicly exposed faculty to an OpenAlex profile. This includes four names and eight incidents for which I have enough information to assign a department and university but not enough to identify the accused individual.

3.2 Identification Strategy and the Construction of Control Research Departments

The empirical strategy examines changes in published research output at two time points: (A) when the incident occurred and (B) when it was reported by the media. The comparison period is the time before the incident occurred, before media reporting, or before the accused faculty member departed the institution, respectively.[8] To ensure that I estimate the causal effect of interest and not some other influence that is correlated with the passage of time, my specifications include time fixed effects. However, there may still be time-related confounders such as policies or event-related trends. To mitigate this threat to identification, I implement a CEM approach, as described below.

To construct the control research departments, I identify all U.S. universities that have been classified as research focused by the Carnegie Classification at least once during the period of interest (1980 to 2024). First, I collect the complete sample of universities that existed in the United States between 1980 and the most recently available year, 2023, using IPEDS data. IPEDS is a set of surveys conducted annually by the U.S. Department of Education's National Center for Education Statistics (NCES). Institutions that participate in a federal student aid program are required to respond by the Higher Education Act of 1965.

After that, I retain universities that have been categorized as research active[9] at least once in the history of the Carnegie Classification. For that, I construct a panel dataset of the Carnegie Classifications between 1974 and 2023. I keep all university names that have ever been designated as a "Research University" during this time period. I then link the

and was likely disconnected from the interpersonal dynamics at the department of interest.

[8] In approximately 65% of the reported cases, the accused faculty member left the institution.

[9] Research-active universities are defined as those listed by the Carnegie Classification in one of the following three categories: "Research 1 - Very High Research Spending and Doctorate Production," "Research 2 - High Research Spending and Doctorate Production," and "Research Colleges and University."

IPEDS database with the selection of research-active universities, which results in a dataset of the universe of research-focused universities between 1980 and 2023. This resulted in 1,152 research-focused universities.

In a last step, I link these universities to OpenAlex. Out of the 1,152 universities, 582 cannot be matched. I manually inspect all of these and find that they are not research-intensive universities, likely explaining why they cannot be linked to OpenAlex. I exclude these universities, which leaves 564 control universities. From these, I randomly select one department for each of the nine disciplines defined by Clarivate's Core Collection. In total, I construct 5,076 control departments (564 control departments per discipline). Because this is the universe of research-active university–discipline pairs, a subset of these university–department pairs is, by definition, treated. Instead of removing them from the control set, I account for that using a staggered treatment design. This is discussed in Section 4.

3.3 Variable Construction

Outcome Variables

Number of Publications per Department Member The publication data are assembled at the department–year–publication level and then aggregated at the department–year or department–year–subgroup (for example, women faculty) level. To capture the productivity of all department members equally, I count within-department co-authored publications multiple times. For example, if two department members co-author together, their publication counts as two publications.

I measure department–year productivity by calculating the number of publications published every year by a given university department divided by the number of department members. This measure has two useful properties. First, the linear transformation of the number of publications in a given department into a per capita value makes it independent of department size and allows for easy comparison across departments. Second, publication count data often follow a power-law distribution with a tail of zeros. This data structure

weights a switch from zero to one more heavily than a shift across nonzero real numbers (Chen and Roth, 2024). Because my dataset is assembled at the university department level, the power-law distribution is less pronounced, as shown in Figure 2. The transformation into a per capita unit removes the concern entirely, as illustrated in Figure 3.

Gender I infer the department members' genders on the basis of their first and, if available, middle names using genderize.io. I use a statistical probability cutoff point of 90% for either the first or middle names or 75% for first and middle names jointly.

Treatment Indicators I collect all publicly available materials on sexual misconduct cases in the U.S. academy (Libarkin, 2024). These include news articles, legal documents, proprietary documents posted on digital document libraries, and information shared in the Academic Sexual Misconduct Database. I have over 1,300 pieces of information for the 359 incidents. Using this rich set of information, I code treatment indicators for *occurrence of a sexual misconduct* and *public reporting of a sexual misconduct incident* at the year level.

Spillover Effects by Gender I investigate whether the impact of sexual misconduct is different for men and women, after inferring their *gender* from their first and, if available, middle name using genderize.io.

Ability to Withdraw from Research Activities I investigate whether the ability to shift away from research modulates the effect of sexual misconduct occurrences on productivity. This is an ex post analysis that allows me to understand whether, in the extreme case, differential productivity effects would be seen. To do so, I define two groups of departments. The first consists of departments where faculty cannot remain employed while stopping research and maintaining the same earnings. These include departments such as biology. The second group consists of departments where faculty can easily remain employed while stopping research and maintaining the same earnings. These include departments of clinical medicine, where one can shift to patient-focused work, and architecture, where one can work with clients on projects.

This measure captures three dimensions of a given department: the availability of alternative tasks (such as patient-focused work), the financial reward for these alternative tasks, and the ease of withdrawing from research activities while remaining on the faculty. I asked two LLMs (a model by Anthropic and a model by OpenAI) to generate categories ranging from "very low" to "high." I then aggregated these scores and assigned a binary variable of either "no easy switch" or "clinical medicine/professional." Although this is admittedly a crude measure of the ability to withdraw from research activities, manual inspection suggests that the measure is sensible. For example, departments in life sciences are categorized as difficult to withdraw from, whereas departments in clinical medicine, law, or architecture are categorized as easy to withdraw from.

Covariates To assign a research discipline to each department, I use the nine primary disciplines from Clarivate's Core Collection (https://mjl.clarivate.com/help-center): Agriculture, Biology & Environmental Sciences; Arts & Humanities; Business Collection; Clinical Medicine; Electronics & Telecommunication Collection; Engineering, Computing & Technology; Life Sciences; Physical, Chemical & Earth Sciences; and Social and Behavioral Sciences.

To determine whether a university is public or private, I use IPEDS' institutional datasets. I can match all 139 unique treated institutions with IPEDS data. The 564 control universities are part of the IPEDS dataset by design.

The research intensity at a given university is assigned using ShanghaiRanking's Global Ranking of Academic Subjects (GRAS), which contains rankings of universities in 55 subjects across natural sciences, engineering, life sciences, medical sciences, and social sciences (Shanghai Consultancy, 2024). These rankings are useful for the purposes of this research because they incorporate research output and quality, faculty quality, and research collaborations.[10] In contrast, the U.S. News Education Ranking focuses on research output and student experience, and the Carnegie Classification rankings are based on the number of doctoral degrees granted and the amount of research expenditure. I assign the GRAS to departments at the

[10]More information about their methodology can be found here: `https://www.shanghairanking.com/methodology/arwu/2024`

university–subject level.

An institution's behavior is likely shaped by the laws under which it operates. To account for that, I construct a categorical variable at the state level that captures the extent to which the state's legal framework deviates from federal law within four categories: discrimination, workplace hostility, sexual harassment, and retaliation. To do so, I hand-collect information about the extent of legal deviations for each category.

Department size is determined by the number of researchers publishing in a given year. The ideal measure for department size would be based on administrative data. In the absence of that, I use the number of active scholars in the department as a proxy for department size. I use this approximation for department size because the sample in this study centers around research-focused universities. Therefore, scholars in these institutions are typically expected to maintain active research agendas as a core component of their responsibilities. This makes the number of publishing researchers a reasonable proxy for overall department size, as non-publishing faculty members are likely to be rare exceptions rather than the norm in such settings. Additionally, this measure captures the intellectually active portion of the department, which is particularly relevant when examining research productivity and scholarly output.

4 Empirical Strategy

To estimate the causal effect of publicly exposed sexual misconduct incidents on academic department outcomes, I implement a staggered DiD design. My approach incorporates CEM as a preprocessing step to improve covariate balance before estimating treatment effects, allowing for causal inference under more credible assumptions. This section discusses empirical challenges, the matching design, and the regression specifications.

Departments experience the two "treatments"—the occurrence of an incident of misconduct and public reporting of it—at different points in time. The treatment effect is identified by comparing outcomes in treated departments before and after exposure, relative to matched

control departments that had not yet been treated or were never treated.

A department is classified as a valid control in a given treatment window if it did not experience a public sexual misconduct case within a symmetric 5-year window around the treatment year. I conduct robustness checks that tighten this requirement in two ways: (i) requiring that no other department in the same university was treated during the window, and (ii) restricting controls to departments in universities that were never treated.

The treatment in my setting is not randomly assigned. A case only becomes public knowledge if some involved parties are willing to file a lawsuit or make a public announcement, or if local or national news media outlets are motivated and able to obtain evidence. It is expected that this is more likely for sexual misconduct cases that are more severe or persistent and those that take place at public universities where information is easier to access, for example, via freedom of information requests. To account for potential differences in observed covariates, I use a multistage CEM approach.

This multistage CEM approach directly addresses several causal challenges. First, the CEM preprocessing helps mitigate bias by ensuring that treatment and control departments are balanced on key covariates that might drive treatment effect heterogeneity. Second, by preprocessing data through CEM before parametric analysis, I reduce the model dependence and sensitivity to specification choices (Ho et al., 2007). The matching procedure creates more comparable treatment and control groups, addressing the fundamental concern that departments experiencing misconduct incidents may systematically differ from those that do not.

Additionally, the inclusion of university–department fixed effects in all specifications provides protection against confounding from institution-level factors. By differencing out all time-invariant university department characteristics, as well as any university-wide shocks that affect all university departments equally, this approach isolates the department-level impact of sexual misconduct incidents. This is particularly important given that universities may differ systematically in their reporting procedures, institutional cultures, and responses to misconduct cases. The university department fixed effects ensure that our estimates reflect

the causal impact of misconduct at the department level rather than capturing heterogeneity at the level of university departments.

I estimate the average treatment effect at the department level. This captures the collective impact on all department members, including both direct effects on those most immediately affected and any spillover effects on their colleagues. The nature of my comparison group leads me to estimate a lower-bound effect when examining the impact of occurrence of sexual misconduct. This is because I compare departments where sexual misconduct took place and became public to a mixture of departments where sexual misconduct did not take place and those where sexual misconduct took place but did not become public. I probe robustness to my results with a set of empirical and statistical tests, as outlined in Section D.

My baseline model is a simple DiD estimator, which I build out into a staggered DiD design (often called two-way fixed effects). Recent methodological developments allow me to address some empirical challenges associated with staggered DiD design (de Chaisemartin and D'Haultfœuille, 2020; Callaway and Sant'Anna, 2021; Goodman-Bacon, 2021; Sun and Abraham, 2021; Baker, Larcker and Wang, 2022). Specifically, standard two-way fixed effects estimators can produce biased estimates in settings with heterogeneous treatment effects over time because they implicitly use already-treated units as controls for newly-treated units. This creates a weighted average of treatment effects in which some effects receive negative weights, potentially leading to estimates with incorrect signs or magnitudes. Because my data have treatment points that span over 40 years, likely creating heterogeneity in treatment effects across cohorts, and I use CEM, my preferred specification uses the Sun and Abraham (2021) estimator. This estimation method handles staggered DiD designs with treatment effect heterogeneity across cohorts and allows for matching weights.

To account for uncertainty across all stages of analysis, I employ cluster-robust standard errors at the university department level, addressing potential correlation between outcomes within departments over time.

4.1 Coarsened Exact Matching

I employ a multistage approach that combines preprocessing through CEM (Iacus, King and Porro, 2011, 2012) with subsequent regression-based analyses. This approach pairs departments that experienced sexual misconduct incidents with similar departments that did not experience such incidents according to prespecified covariates. One of the key advantages of this approach is that it reduces model dependence by first addressing selection bias through matching before conducting parametric analysis (Ho et al., 2007). For simplicity, I use *publication* as the outcome variable Y from here on.

In the first stage, I implement CEM to create appropriate comparison groups by matching departments that experienced sexual misconduct incidents with similar departments that did not. To satisfy the conditional independence assumption, I match on covariates that simultaneously predict both departmental research activity and the likelihood of experiencing sexual misconduct incidents or public reporting of them. Specifically, I match on research discipline, the legal framework at the state level, department size, and research intensity. These covariates block the main confounding pathways between misconduct occurrence and productivity. For example, larger departments have both more opportunities for misconduct incidents and greater research output By matching on these variables, I ensure that treated and control departments differ primarily in their treatment status rather than in other characteristics that could confound the relationship.

Research Discipline. Research disciplines strongly influence publication patterns because of field-specific characteristics and norms. Thus, I stratify the matching process by research discipline. I use Clarivate's nine primary research categories defined in its Core Collection; each of the 142 disciplines in the treated dataset is linked to one primary research discipline.

Legal Framework at the State Level. An institution's behavior is likely influenced by the laws that govern it. To account for that, I construct a categorical variable at the state level that captures deviations of the state-level legal framework from federal law related to

discrimination, workplace hostility, sexual harassment, and retaliation. To do so, I collect legal information and classify each state by the extent to which it deviates from federal law. I produce three coarsened strata: states in line with federal law, states that provide minor expansions compared with federal law, and states that provide wider or significant expansions compared with federal law.

Research Intensity. The research intensity at a given university is assigned using the ShanghaiRanking's GRAS, which contains rankings of universities in 55 subjects across natural sciences, engineering, life sciences, medical sciences, and social sciences (Shanghai Consultancy, 2024).

4.2 Staggered Difference-in-Differences Estimation

Analysis 1: Overall Treatment Effect Analysis Using the matched sample from Stage 1, I estimate the overall effect of sexual misconduct incidents on department productivity. I apply the following weighted regression model to the matched data:

$$Y_{it} = \alpha_i + \lambda_t + \beta_1 Post_{it} + \tau(Treat_i \times Post_{it}) + \gamma DeptSize_{i,t-7} + \varepsilon_{it} \tag{1}$$

where Y_{it} is the number of publications for university department i at time t, α_i are university department fixed effects, λ_t are time fixed effects, $Treat_i$ is a binary indicator equal to 1 for departments that experienced a misconduct incident (and 0 otherwise), $Post_{it}$ is a binary indicator equal to 1 for periods t after department i experienced a misconduct incident or the public reporting of it (and 0 otherwise), τ is the coefficient of interest representing the causal effect of sexual misconduct exposure on departmental productivity, γ controls for lagged department size, and ε_{it} is the error term.

Standard errors are clustered at the university–department level, and control departments

are weighted using the CEM stratum-specific weights:

$$w_i = \begin{cases} 1 & \text{if } i \text{ is treated} \\ \frac{m_T^s}{m_C^s} & \text{if } i \text{ is a control unit in stratum } s \end{cases} \tag{2}$$

where m_T^s is the number of treated units in stratum s, and m_C^s is the number of control units in stratum s.

I report the average treatment effect over the five post-treatment periods, for which I take the simple average of τ,

$$\bar{\tau} = \frac{\sum_{t=1}^{5} \tau_{it}}{5} \tag{3}$$

with a standard error of the mean of the post-treatment coefficients. I then report the full behavior in standard event-study plots.

Analysis 2: Heterogeneity Analysis by Gender Continuing with the matched sample, I investigate how the impact of misconduct incidents varies between men and women within departments. For this subgroup analysis, I estimate the following:

$$Y_{it}^g = \alpha_i^g + \lambda_t^g + \beta_2^g Post_{it} + \tau^g (Treat_i \times Post_{it}) + \gamma^g DeptSize_{i,t-7} + \varepsilon_{it}^g \tag{4}$$

where Y_{it}^g represents the productivity measure for subgroup g in university department i at time t, τ^g captures the causal effect of sexual misconduct exposure specific to subgroup g, and γ^g controls for department size.

4.3 Staggered Event Study Design

To investigate the temporal dynamics of the treatment effect, I estimate an event-study specification:

$$Y_{it} = \alpha_i + \lambda_t + \sum_{k \neq -1} \beta_k \cdot \mathbb{1}(r_{it} = k) + \varepsilon_{it} \tag{5}$$

where $r_{it} = t - T_i$ is time relative to the treatment year. The year prior to the treatment

is the omitted category ($k = -1$). I plot the coefficients β_k to visually assess pre-trends and time-dynamic treatment effects. My preferred specification uses the Sun and Abraham (2021) estimator because it allows for matching weights and takes into account heterogeneous treatment effects over time.

5 Results

5.1 Descriptive Statistics

To start, I describe the sexual misconduct cases that took place in research-focused universities, were committed by faculty, and became public. These misconduct cases occurred in all academic disciplines, from arts and humanities (34.8% of all cases) and social and behavioral Sciences (26.2%) to engineering, computer and technology (4.2%) and business (1.1%). The cases also differed in severity. Out of all 359 misconduct cases, 132 involved sexual assault, 228 involved sexual harassment, 55 involved sexual advances, and 101 involved sexualized comments. These categories are not mutually exclusive; several cases involved multiple types of sexual misconduct. These instances of misconduct were mostly perpetrated by tenured faculty (in 67% of the cases).[11] In over half the cases, there was a repeat offender. Specifically, in 12.5% of the cases, the accused faculty faced two allegations, and in 39.8% of the cases, they faced three or more.

—— * ——

About here:

Table 1: Characteristics of Sexual Misconduct Cases in University Departments

—— * ——

The search of publications for the 5,076 universities from 1980 until 2024 yielded a total of 1,374,089 publications and 7,820,515 authors. Gender could be assigned to 80.26% of the authors. Of all authors, 56.49% are algorithmically determined to be men and 23.78% women.

[11] It is likely that this number is slightly biased upward because in some cases, it was unclear whether the stated rank was at the time of misconduct or the time of reporting.

The remaining authors (19.74%) either carry a gender-ambiguous name or have a name that is unknown to the database. This includes authors for whom I only have the initials.

Table 2 shows the summary statistics of the treated and control departments in the year used for matching. After matching, there are 201 departments in the treatment group and 1,982 in the control group. The two groups look descriptively and unconditionally similar when we compare their legal environment. What is notable, however, is that the group of treated departments has a higher share of public universities (73.63% vs. 46.92%), a larger average number of publications (61 vs. 18), and a larger average number of department members (90 vs. 31).

___ * ___

About here:

Table 2: Summary Statistics of Treated and Control Departments

___ * ___

The CEM reduces the differences between the treated and matched control departments noticably from a multivariate L1 distance of 0.71 to 0.5, as shown in Table B2. The L1 distance is reduced by 54% for *discipline* and to almost zero for the type of institution (public or private), the matching based on state laws for hostility and harassment, and the Shanghai ranking. The L1 distance for department size is reduced by 67%, although some imbalance remains.

Table B1 shows that control departments have higher baseline publication rates than treated departments, both before (2.18 vs. 1.53) and after matching (2.98 vs. 1.51). This persistent difference in outcome levels does not compromise identification because the matched control departments serve as valid counterfactuals for treated departments despite different productivity levels. The identifying assumption is that matched departments would follow parallel trends in the absence of treatment, not that they have identical baseline outcomes. The DiD design with department fixed effects eliminates time-invariant level differences, identifying the causal effect through within-department variation in productivity following incidents of misconduct.

5.2 Occurrence Shows No Discernible Overall Effect

Next, I turn to the main results. Recall that the analysis simultaneously examines the effect of occurrences of sexual misconduct and the effect of public reporting on scientific production. Figure 4 shows an event-study plot of the former. The overall finding is that an occurrence of sexual misconduct has no discernible effect on the average number of publications per department member. Column 1 uses the full sample. These cases were publicly reported at a later date. The average time that passed between the occurrence and public reporting is 7 years and the median is 3 years; the full distribution is shown in Figure 5a. Thus, these results include any effects on productivity of the public disclosure of sexual misconduct. To probe the robustness of my results, I remove all incidents for which public reporting occurred within the five-year window of this analysis. This subsample analysis presents a trade-off. Although it yields a cleaner estimation because there were no public reporting events within five years, it is likely selecting on a type or degree of misconduct. Sexual misconduct incidents that take longer to become public may be more complex and thus require a longer time window for the investigation. They may also be more severe. The result is robust to this subsample, as presented in Figure 5b.

___ * ___

About here:

Figure 4: Effect of Occurrence of Sexual Misconduct on the Average Number of Publications per Department Member

___ * ___

Figure 5: Effect of Occurrence of Sexual Misconduct on the Average Number of Publications per Department Member (Restricted to Cases That Were Not Publicly Reported Within Five Years)

___ * ___

Estimates of the effect of an occurrence of sexual misconduct on productivity may be attenuated toward zero by contamination in the control group: departments that experienced unreported misconduct would dilute any true negative effects. Therefore, there is a possibility that the true effect may be non-negligible. I address this through both empirical and statistical

approaches.[12] First, I create a control group in which one would expect a lower level of contamination by restricting it to departments with above-median representation of women for a given discipline at time $t-1$, because research shows that higher representation of women is associated with a decreased likelihood of sexual misconduct occurring (Mansfield et al., 1991; Welsh, 1999; Folke and Rickne, 2025; Subramani and Gorbatai, 2025). Figure 6 shows that even with this refined control group, I find no discernible effect of sexual misconduct occurrence on departmental productivity.[13]

Second, I conduct statistical exercises to assess the plausibility of meaningful negative effects being masked by contamination. This approach is in the spirit of Manski (2003), Altonji, Elder and Taber (2005), and Oster (2019), who consider how selection into control or treatment affects the stability of an observed treatment effect. I calculate the contamination levels required to reconcile the observed effect of 0.022 with a meaningfully negative true effect. I define a meaningful negative effect as -0.056 publications per faculty-year, which represents one full publication per year in a median-sized department of 18 faculty members.[14] Using the contamination adjustment formula, True Effect $=$ Observed Effect$/(1-c)$, where c represents the contamination rate, I find that even accounting for statistical uncertainty, contamination would need to exceed 59% of control departments for such an effect to be plausible. This would imply a total prevalence of sexual misconduct across all departments of at least 64%, which far exceeds the value of 38% documented by Boudreau et al. (2023).

Beyond this contamination bounds analysis, I conduct additional robustness tests examining four scenarios that vary both contamination levels (10% vs. 29%) and true effect sizes (negligible vs. meaningfully negative). These exercises, detailed in Appendix D, consistently support the interpretation that the observed null effect likely represents a genuine absence of productivity impacts. Although I cannot definitively rule out all alternative explanations, the convergence of evidence—the null finding with a refined control group, the need for implausibly high contamination rates for meaningful effects, modest adjusted effects under realistic con-

[12] This is not preregistered.

[13] Although this is still a null effect, the sign flips from positive to negative. This is in line with my prior that the effect may be negative but likely limited, as outlined in Section 2.

[14] My rationale is that one publication is large enough to show up in the statistics of a department.

tamination, and statistical rejection of scenarios with meaningful negative effects—suggests that sexual misconduct has negligible effects on departmental research productivity at the occurrence stage.

5.2.1 Effects Are Similar for Women and Men

Most victims of sexual misconduct in academia are women. Informal information flow via whisper networks may suggest that there are sorting and disruptions that may not gender-neutral. Specifically, it may be possible that women adjust their work patterns in some form, for example, via collaboration networks, as suggested by Gertsberg (2022). This may have downstream consequences on their productivity. Therefore, Columns 3 and 4 of Table B3 report the results split by gender. The estimates are comparable, and I do not observe a noticeable difference among women (nor men).

5.2.2 Withdrawal in Departments Where Switching Away from Research Is Easy

There may be particular instances in which productivity is dampened to a notable extent, however. For example, some organizational regimes allow for faculty to shift away from their core activity of research Cortina and Berdahl (2008). As an exploratory analysis,[15] I analyze whether withdrawal from research activities can be observed in departments where switching to non-research activities is easy and costless. To explore this possibility, I split the sample into two groups. The first consists of departments where switching is easy and relatively costless. This includes departments in clinical medicine and professional fields such as architecture, law, physical therapy, and veterinary sciences. The second group consists of departments where it is difficult to remain on the faculty and switch away from research activities. Examples include biology, astronomy, chemistry, and education.

Figure 7 presents the results. Panel 7a reports the event-study estimates for departments where switching away from research is easy. Panel 7b reports the event-study estimates for departments where a switch away from research is relatively difficult. Although there is no

[15]This analysis is not preregistered.

discernible effect on the research output for departments where it is harder to move away from research, I document significant withdrawal in departments where withdrawal is easier. On average, a department member in this category reports 0.12 fewer publications for the five years after sexual misconduct occurs; the p-value is below 0.01. As Panel 7a shows, the effect size increases somewhat over time, starting at -0.07 in the year of the event and reaching -0.18 in the fifth year after the event.

___ * ___

About here:

Figure 7: Comparison of Departments with Varying Degrees of Ability to Switch

___ * ___

The main finding of this first set of analyses is a null effect: the occurrence of sexual misconduct has no discernible impact on average departmental publication rates, whether examining all faculty, women, or men. Appendix D probes confidence of this null effect, addressing potential contamination in the control group and confirming that the negligible productivity effects are genuine. However, there is a limited set of departments—perhaps a form of edge case—where I do find a decrease in productivity. This occurs specifically in departments where it is relatively easy for faculty to shift away from research activities, such as clinical medicine and other professional fields, where affected departments show a decline in the five years following the occurrence of sexual misconduct.

5.3 Public Reporting Reduces Average Number of Publications

Next, I turn to a set of analyses of the effect of public reporting of sexual misconduct on university department productivity. Figure 8 shows an event-study plot, which demonstrates that the public reporting of sexual misconduct has a negative effect on the average number of publications per department member. My preferred estimate combines CEM weights with cohort weights to adjust for heterogeneous treatment effects. It shows that sexual misconduct has an overall negative effect on the number of publications per department member, ranging from -0.04 to -0.13 publications per year. In other words, over the course of five years after a

misconduct case becomes public, for the median department of 18 faculty members, it resulted in nine fewer publications than in comparison departments. Figure B5 shows that this result is robust to alternative specification.

—— * ——

About here:

Figure 8: Effect of Public Reporting of Sexual Misconduct on the Average Number of Publications per Department Member

—— * ——

The result is robust to various approaches to DiD design, as shown in Columns 1–5 of Table B5. Specifically, I start with an analysis in the form of a simple DiD design. Column 1 reports the average treatment effect over the five years following the public reporting of a sexual misconduct incident. Column 2 reports the same estimates but in the form of an event study. The average annual treatment effect over the five years after the public reporting of sexual misconduct is -0.04 publications per department member. The p-value is below the common α threshold of 0.05. This compares to an annual treatment effect ranging from -0.01 to -0.07 in the simple event study in Column 2.

The estimates based on the matched sample are qualitatively similar but noticably larger. Column 3 reports the average treatment effect over the five years following the public reporting of a sexual misconduct incident. Column 4 reports the same estimates but in the form of an event study. Column 5 reports the matched event study using the Sun and Abraham (2021) estimator. The average annual treatment effect over the five years after the public reporting of sexual misconduct is -0.1 publications per department member. This estimate is directionally significant. This compares to an annual treatment effect ranging from -0.05 to -0.14 and from -0.04 to -0.13, respectively, in the event-study estimates in Columns 4 and 5 of Table B5. The estimates from the event study suggest that the productivity effect appears over time, with the greatest decrease in productivity in year three.

5.3.1 Public Reporting Affects Men More Than Women

The effect varies somewhat between men and women. Table B6 shows that men who were members of a department where sexual misconduct occurred and was publicly reported published 0.11 ($p < 0.10$) fewer publications per year, on average, than men in comparable departments without a publicized sexual misconduct case. For women, the estimated effect is -0.06 ($p < 0.05$).

These estimates should be interpreted as intention-to-treat effects that capture department-wide impacts rather than individual-level consequences for directly affected parties. This intention-to-treat analysis indicates that sexual misconduct incidents create organizational externalities that extend beyond victims and perpetrators, affecting the productivity of entire academic departments. Although individual victims undoubtedly experience severe personal and professional consequences, my results demonstrate that organizational costs are substantially broader, providing an economic rationale for institutional investment in prevention and response systems.

6 Discussion and Conclusion

In this paper, I ask, *how do sexual misconduct incidents in university departments affect scientific production?* I build a panel dataset with over 5,000 university departments across nine disciplines between 1980 and 2024. The treated departments consist of a comprehensive set of departments at research-focused universities in the United States with sexual misconduct cases that were publicized within that time period and the accused individual had the role of a faculty member. To construct a set of comparable control departments, I use a list of unique research-focused universities based on IPEDS data and the Carnegie Classification. I then randomly select nine departments per university as possible control departments, making 5,076 in total.

The results of the preregistered analysis indicate that *public reporting* of sexual misconduct incidents decreases university departments' overall productivity but that the *occurrence*

of sexual misconduct does not. Exploratory analysis shows that the latter non-discernible effect masks some heterogeneity. I find that in departments where withdrawal from research is possible and costless—for example, in clinical medicine, by switching to patient-focused activities—sorting does occur and department members shift away from research activities.

These findings are consistent with stigma-based mechanisms that lead to a distinction between contained and spillover effects. Directly affected individuals may avoid discussing the stigmatized event with others because of self-blame, shame, and fear of negative judgment, which contains knowledge of the misconduct among those with direct involvement or close ties. In addition, reporting and resolution processes are typically designed to prioritize confidentiality. However, when sexual misconduct becomes public, different mechanisms emerge. Public reporting creates reputational damage that extends beyond direct victims through stigma by association, as individuals with observable ties to the accused—whether through department membership or shared group characteristics—experience productivity losses even when not directly involved in the misconduct itself.

6.1 Contributions to the literature

This paper contributes to several strands of the literature. First, a robust body of research examines how workplace sexual misconduct affects victims and other directly involved individuals (Schneider, Swan and Fitzgerald, 1997; Sims, Drasgow and Fitzgerald, 2005; Cortina and Berdahl, 2008; McLaughlin, Uggen and Blackstone, 2017; Adams-Prassl et al., 2024). This study extends that important work by shifting the focus to the effects on organizations. My findings suggest that sexual misconduct incidents do not have a direct effect on organizations, except for a shift away from research to other activities in professional fields where that is a realistic option. However, the public reporting of sexual misconduct leads to reputational damage and a decline in research publications.

Second, this study contributes to the literature on organizational misconduct. Past work examines how public revelations of sexual misconduct affect organizations (Yenkey, 2018; Stroube and Zavyalova, 2024; Yenkey and Palmer, 2025). I extend this work by examining

the resulting organizational dynamics when sexual misconduct occurs and when it is publicly revealed. My results suggest that the occurrence and public reporting of sexual misconduct have different effects on organizations. I theorize that the occurrence of sexual misconduct triggers sorting mechanisms, as organization members reallocate their efforts, whereas public reporting activates cognitive processes and reputational management strategies. By examining misconduct at both time points, this study provides a more nuanced understanding of the organizational sociology and economics of misconduct. Moreover, by demonstrating that organizational responses operate through fundamentally different mechanisms at the time of occurrence and the time of public disclosure, this study opens new avenues for understanding how organizations respond to internal crises more broadly. The findings suggest that theories focusing solely on post-revelation effects may miss critical organizational dynamics that shape long-term outcomes.

Third, this work extends the literature on the science of science. An emerging stream of research examines how increased societal awareness of sexual misconduct affects scientific production. Recent work shows that the #MeToo movement altered collaboration structures for women (Gertsberg, 2022) and that harassment accusations affect collaboration patterns and recognition of accused faculty members (Maimone et al., 2025; Widmann, Rose and Chugunova, 2022; Luo and Zhang, 2022). I advance this conversation by shifting the analytical focus from individual-level effects to the impacts on university departments, examining how sexual misconduct incidents reverberate throughout entire academic units using a novel methodological approach that captures effects both at the time of incident occurrence and upon public revelation. I find that, when sexual misconduct *occurs*, departments where the shift away from research is easy and costless see a reduction in scientific publishing. However, public reports of sexual misconduct have a negative effect on university departments' research productivity regardless of the type of department. The findings indicate that factors traditionally considered peripheral to scientific productivity, such as organizational climate, interpersonal dynamics, and workplace disruptions, may play more central roles in shaping research outcomes than previously recognized.

6.2 Implications

The results presented in this paper can inform practice and policy. They show that organizational costs arise specifically from public disclosure rather than from misconduct itself. This distinction between the effects of occurrence and disclosure suggests that organizations may face perverse incentives to keep misconduct hidden, to the detriment of victims.

This perverse incentive structure helps explain why organizations often appear reluctant to provide full transparency: they must balance moral imperatives against significant organizational consequences. As social pressure increasingly demands transparency, sexual misconduct is transforming from a human resource concern into a strategic organizational challenge.

For policymakers, this implies that protecting victims and maintaining organizational productivity may require differentiated policy approaches as organizations navigate competing demands from legal frameworks, ethical obligations, and performance concerns.

6.3 Misconduct More Broadly

A question arising from my findings is whether they generalize to other forms of organizational misconduct, such as financial fraud (Kang, 2008; Stuart and Wang, 2016; Wang, Stuart and Li, 2021), doping (Yenkey and Palmer, 2025), or scientific fraud (Furman, Jensen and Murray, 2012; Lu et al., 2013; Azoulay et al., 2015; Gross, 2016; Azoulay, Bonatti and Krieger, 2017). There are reasons to believe that these types of misconduct are similar and comparable. For example, when they become public knowledge, they often operate as a scandal and inflict stigma on organizations (Hudson, 2008). However, as Greve, Palmer and Pozner (2010, p. 54) note, "acts are labeled as misconduct whenever they are harmful or morally objectionable," suggesting that there may be more nuance to organizational misconduct than currently acknowledged and that further evidence may be needed to answer this question.

For example, there are reasons to believe that misconduct may relate to organizational performance and reward structures in distinct ways. Both financial and scientific fraud, for example, can be understood in part as corrupted forms of performance enhancement. Although

they are associated with an apparent increase in organizational performance, they exploit legitimate organizational activities to secure personal gains like bonuses, promotions, funding, or status. However, these forms of fraud may differ in their systemic impacts: financial fraud might produce more contained, short-term effects, whereas scientific fraud can redirect entire research fields and alter the direction of scientific inquiry for decades.

Sexual misconduct might operate through somewhat different mechanisms. Rather than representing a corrupted version of normal organizational processes, sexual misconduct involves the abuse of organizational power structures to inflict direct interpersonal harm. That is, the offender is not simply transgressing in core organizational tasks: they are actively exploiting institutionally legitimized hierarchies and authority as tools to cause harm to others. The benefits to perpetrators may be unrelated or counterproductive to organizational performance metrics.

Thus, sexual misconduct may be conceptually different from other forms of organizational misconduct, despite sharing certain organizational consequences such as creating scandals and reputational damage. Although all forms of misconduct violate organizational norms, the underlying mechanisms, motivational structures, and relationships to organizational performance may vary substantially between distinct types of misconduct. Future research should systematically examine these differences to determine whether misconduct types require tailored theoretical and practical approaches.

6.4 Limitations and Future Research

As with all empirical work, this study operates within certain data constraints. The sexual misconduct cases examined here are those that eventually became public, creating a sample conditional on later disclosure. Although a complete census of all incidents would be ideal, this study represents an important first step in understanding the temporal dynamics of organizational misconduct. By demonstrating distinct effects at occurrence and reporting even within this constrained sample, the findings establish both the theoretical importance and empirical feasibility of examining these dynamics. Future research can build on this framework as data

systems evolve to capture a broader range of incidents.

Another possible limitation is that the intent-to-treat nature of my estimates may mask important heterogeneity in individual-level effects. I capture average department-level effects; the consequences for direct victims, accused faculty, and bystanders likely vary substantially. Future research with access to individual-level data could decompose these department-wide effects to better understand the distribution of impacts across different groups within affected organizations.

Several avenues warrant future investigation to refine our understanding of organizational misconduct. Systematic literature reviews or empirical comparisons of conceptually distinct types of misconduct in similar contexts would provide useful insights into the differences between forms of organizational misconduct. Moreover, although this paper advances our understanding of spillovers associated with sexual misconduct, the boundaries of these effects remain unclear. Particularly important is understanding how sexual misconduct affects junior academics and their career trajectories, as well as considering its explanatory power for the disparities in representation of genders seen in certain fields.

Another important direction for future work concerns organizational responses to sexual misconduct. In particular, why do organizations often retain perpetrators rather than terminate them? Adams-Prassl et al. (2024) find that women supervisors are more likely to dismiss perpetrators, and Collis and Van Effenterre (2024) provide exploratory evidence that individuals averse to hostile environments preferentially sort into departments with women supervisors. Understanding these dynamics could inform more effective organizational policies for addressing sexual misconduct while minimizing productivity losses.

Lastly, we have a surprisingly sparse understanding of the processes that occur within an organization following an incident of sexual misconduct. Scholars and policymakers alike would greatly benefit from a richer understanding of these processes and the consequences that the occurrence and processing of sexual misconduct has on organizations.

6.5 Conclusion

This study demonstrates that sexual misconduct creates organizational costs that extend far beyond direct victims. However, these costs arise from disclosure, not occurrence. This asymmetry poses a fundamental challenge for organizations navigating competing demands from ethical obligations, legal frameworks, and performance concerns. Understanding these dynamics is essential for developing policies that protect individuals without inadvertently incentivizing concealment.

7 Figures and Tables

Figures

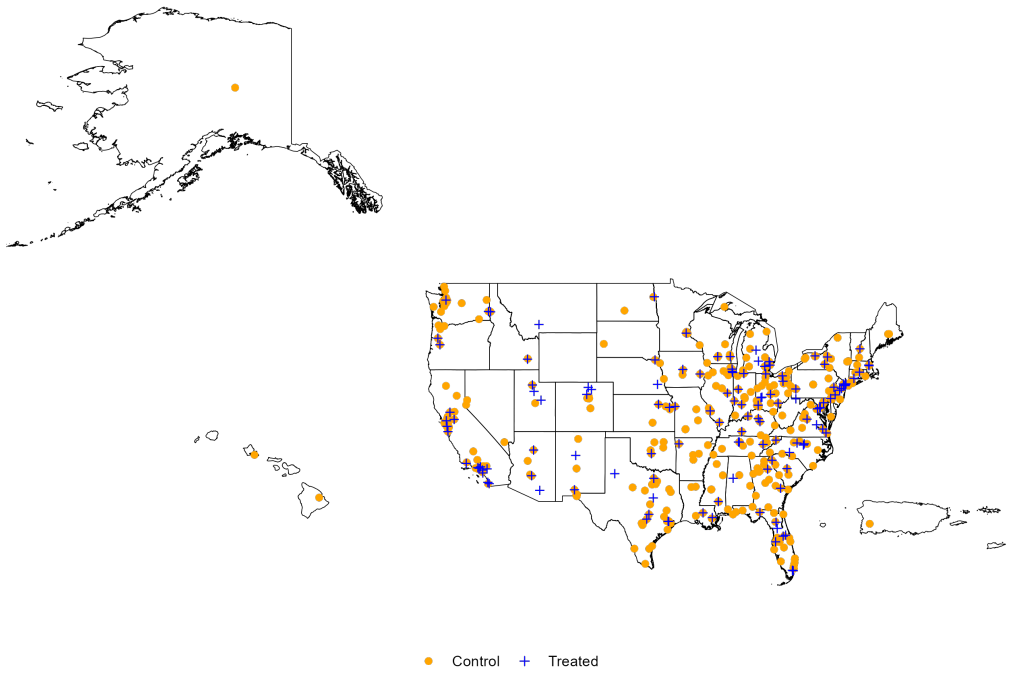
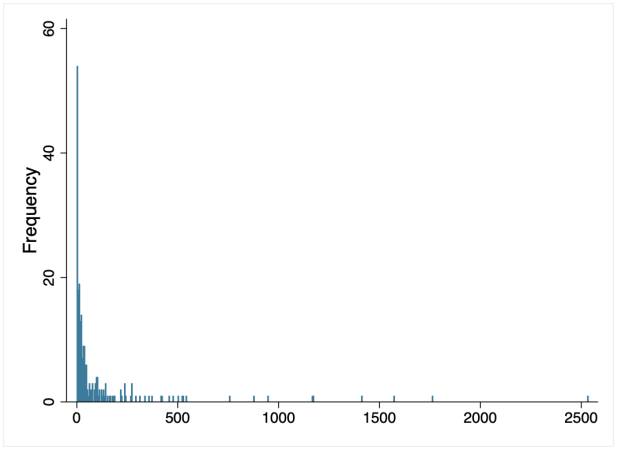
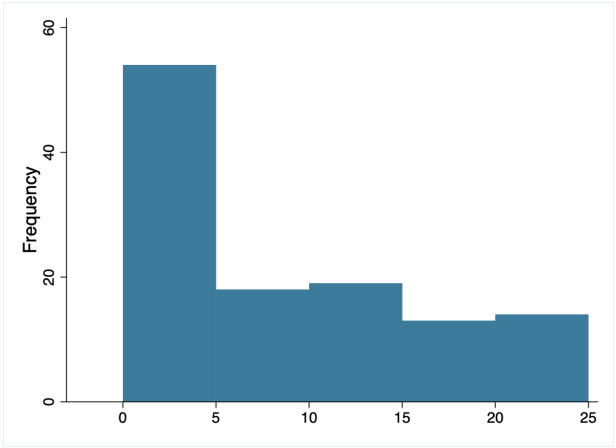


Figure 1: Geographic Distribution of Treated and Control Universities



(a) Distribution of publication counts of treated departments



(b) Distribution of publication counts of treated departments, restricted to the lower 50 percentiles

Figure 2: Histogram of Publication Counts of Treated Departments at Time $t-1$ Relative to an Occurrence of Sexual Misconduct

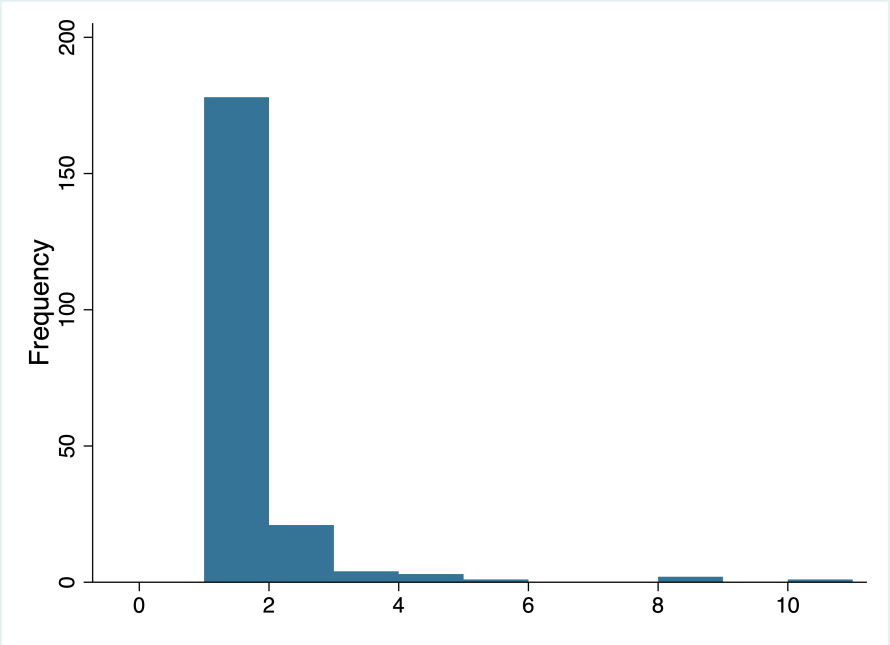


Figure 3: Distribution of the Average Number of Publications per Department Member at Time $t-1$ Relative to an Occurrence of Sexual Misconduct

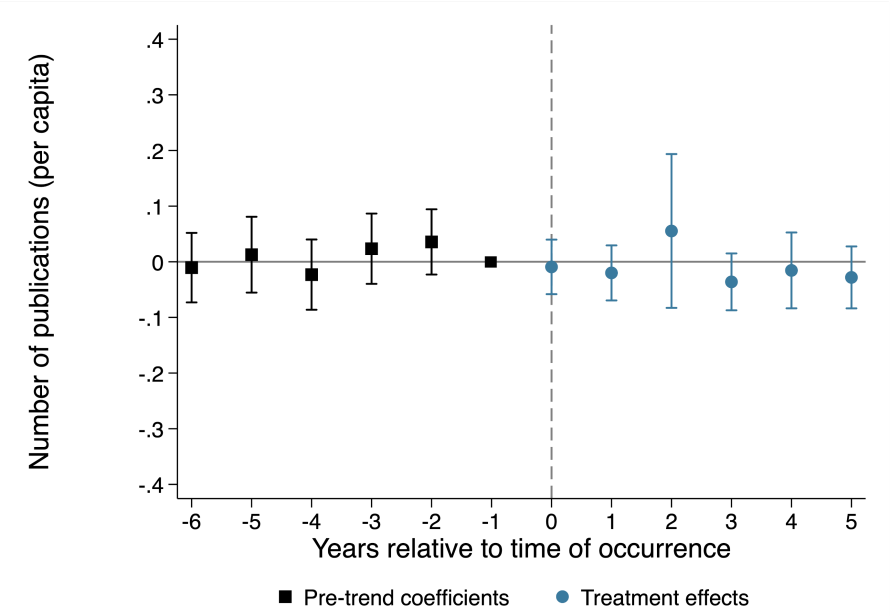
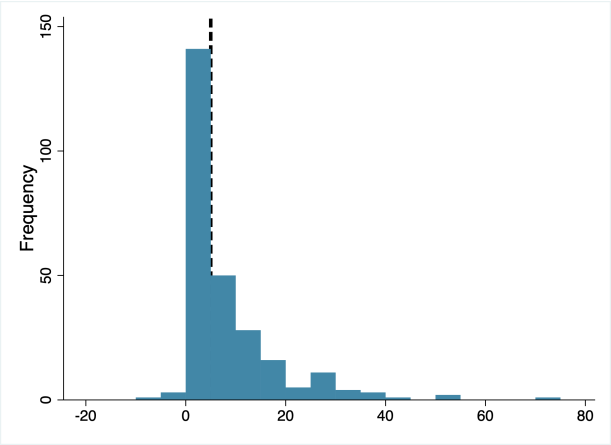
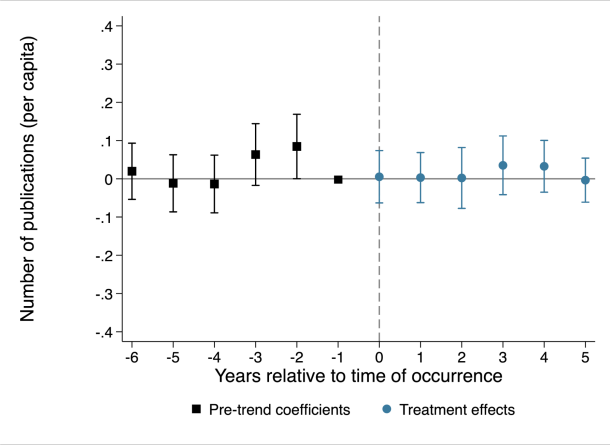


Figure 4: Effect of Occurrence of Sexual Misconduct on the Average Number of Publications per Department Member

Note: The full data can be found in Table B3.



(a) Number of years passed between occurrence and public reporting



(b) Estimation based on cases with more than a 5-year delay in public reporting

Figure 5: Effect of Occurrence of Sexual Misconduct on the Average Number of Publications per Department Member (Restricted to Cases That Were Not Publicly Reported Within Five Years)

Note: The full data can be found in Table B3.

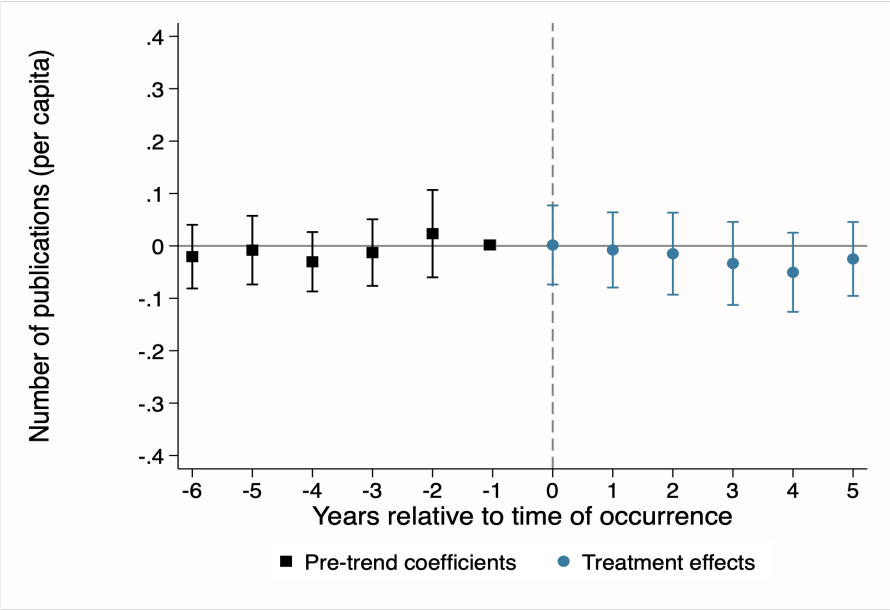
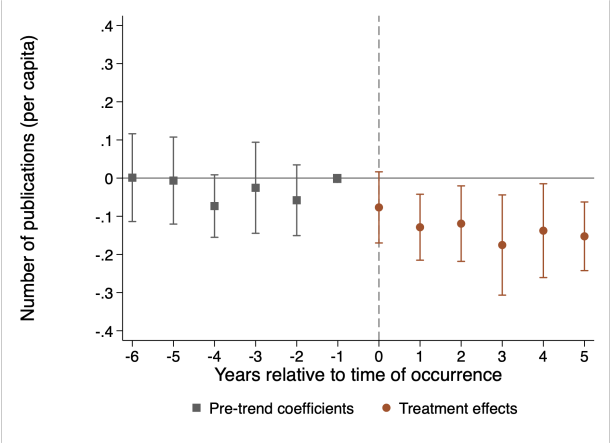
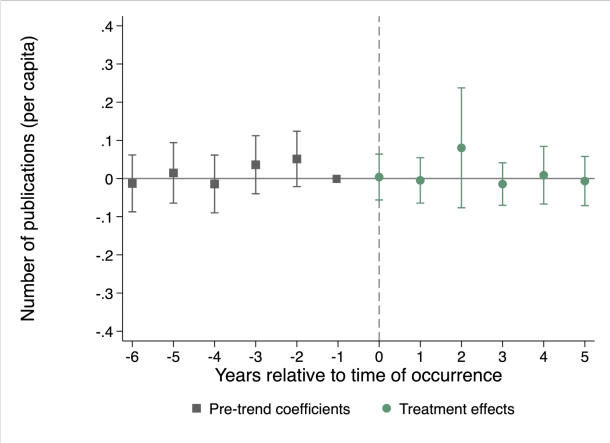


Figure 6: Effect of Occurrence of Sexual Misconduct on the Average Number of Publications per Department Member (Control Restricted to Departments with Representation of Women Above the Median)

Note: The full data can be found in Table B3.



(a) Departments where it is easy to withdraw from research activities



(b) Departments where it is difficult to withdraw from research activities

Figure 7: Comparison of Departments with Varying Degrees of Ability to Switch

Note: The full data can be found in Table B4.

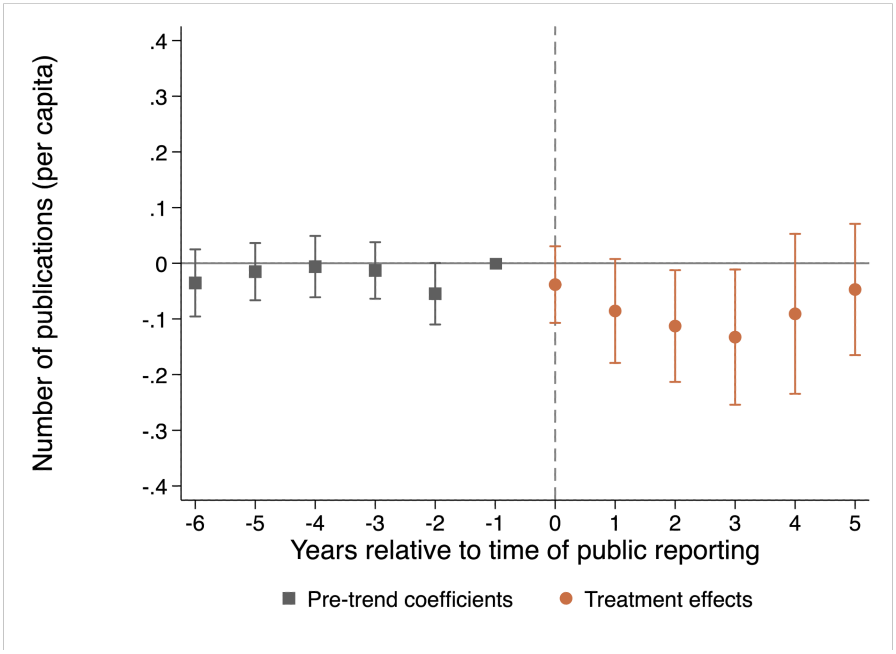


Figure 8: Effect of Public Reporting of Sexual Misconduct on the Average Number of Publications per Department Member

Note: The full data can be found in Table B5.

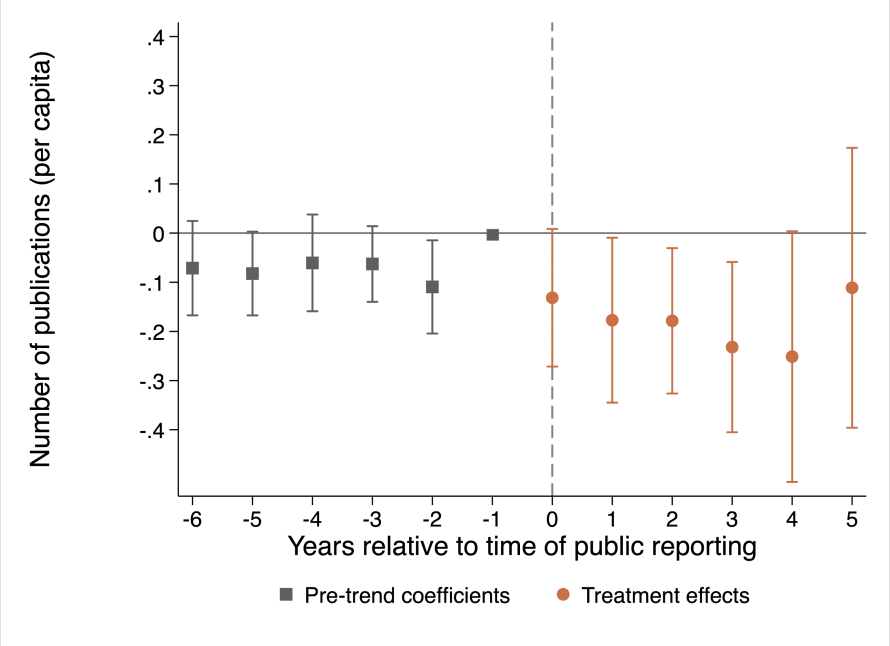


Figure 9: Effect of Public Reporting of Sexual Misconduct on the Average Number of Publications per Department Member (Restricted to Cases That Were Not Publicly Reported Within Five Years)

Note: The full data can be found in Table B5.

Tables

Table 1: Characteristics of Sexual Misconduct Cases in University Departments

	Mean	Median	SD	Min.	Max.
Year of occurrence	2008	2010	9.65	1950	2025
Year of public reporting	2014	2017	7.27	1995	2025
Years of delay	6.96	3	8.80	0	74
Disciplines					
Agriculture, Biology & Environmental Sciences	7.26%	0	0.26	0	1
Arts & Humanities	31.68%	0	0.46	0	1
Business Collection	1.15%	0	0.11	0	1
Clinical Medicine	9.54%	0	0.29	0	1
Electronics & Telecommunication Collection	0.38%	0	0.06	0	1
Engineering, Computing & Technology	3.44%	0	0.18	0	1
Life Sciences	7.6%	0	0.27	0	1
Physical, Chemical & Earth Sciences	9.54%	0	0.29	0	1
Social And Behavioral Sciences	29.01%	0	0.45	0	1
Role/Position of Accused					
Assistant Professor	7.25%	0	0.26	0	1
Associate/Established/Tenured Professor	40.46%	0	0.49	0	1
Full Professor	29.39%	0	0.46	0	1
Emeritus	3.05%	0	0.17	0	1
Not Specified	19.85%	0	0.40	0	1
Type of misconduct					
Abuse of Power & Professional Misconduct	5.34%	0	0.23	0	1
Hostile Environment	5.73%	0	0.23	0	1
Retaliation	4.12%	0	0.20	0	1
Inappropriate Relationship	6.49%	0	0.25	0	1
Sexualized/Inappropriate Comments	3.82%	0	0.19	0	1
Sexual Harassment & Misconduct	26.71%	0	0.44	0	1
Stalking & Intimidation	5.34%	0	0.23	0	1
Sexual Violence & Assault	11.07%	0	0.31	0	1
Substance-related Harassment & Assault	0.76%	0	0.09	0	1
Number of allegations					
Once	35.88%	0	0.48	0	1
Twice	12.60%	0	0.33	0	1
Serial	41.22%	0	0.49	0	1
Unknown	10.31%	0	0.31	0	1
Role/Position of victim					
Peer (Faculty, Researcher, Colleague)	8.02%	0	0.27	0	1
Graduate or Postdoctoral Student	26.34%	0	0.44	0	1
Research or Teaching Assistant	2.29%	0	0.15	0	1
Undergraduate Student	3.05%	0	0.17	0	1
Student (level unknown)	44.27%	0	0.51	0	1
Other/Unknown	16.03%	0	0.37	0	1

Notes: Sample consists of 262 sexual misconduct cases at US research universities that became publicly known between 1980 and 2024 and could be matched to control departments. Years of delay represents the time elapsed between the occurrence of misconduct and its public reporting. Type of misconduct and role/position of victim categories are not mutually exclusive, as cases may involve multiple types and victims. Abuse of power & professional misconduct, hostile environment, and retaliation are coded for context but are not counted as sexual misconduct for the main analysis in this paper.

Table 2: Summary Statistics of Treated and Control Departments

	Mean	Median	SD	Min.	Max.
Control Departments (Never-treated)					
University Characteristics (N = 1,982)					
Public Universities	46.92%	0	0.50	0	1
State Law compared to Federal Law					
Compliant	43.90%	0	0.50	0	1
Slightly Expanded	23.36%	0	0.42	0	1
Significantly Expanded	32.04%	0	0.47	0	1
Department Characteristics (N = 1,038)					
Total publications	18.09	3	49.88	0	969
Articles Share	0.68	0.95	0.43	0	1
Department member count	30.61	1	5	0	2,129
Female member count	8.08	1	26.68	0	565
Male member count	16.71	3	54.42	0	1,178
Treated Departments					
University Characteristics (N = 201)					
Public Universities	73.63%	1	0.44	0	1
State Law compared to Federal Law					
Compliant	46.27%	0	0.50	0	1
Slightly Expanded	13.93%	0	0.35	0	1
Significantly Expanded	38.81%	0	0.49	0	1
Department Characteristics (N = 180)					
Total publications	60.96	26.5	98.53	0	742
Articles Share	0.89	0.97	0.21	0	1
Department member count	89.52	38.5	131.10	0	743
Female member count	26.91	13	41.90	0	313
Male member count	45.82	18	68.11	0	351

Notes: This table presents summary statistics for treated departments (those experiencing publicly reported sexual misconduct) and matched control departments. University characteristics are stable over time. Department characteristics are measured at baseline (t-1).

References

Acquadro Maran, Daniela, Antonella Varetto, and Cristina Civilotti. 2022. "Sexual Harassment in the Workplace: Consequences and Perceived Self-Efficacy in Women and Men Witnesses and Non-Witnesses." *Behavioral sciences*, 12(9).

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

Adut, Ari. 2005. "A theory of scandal: Victorians, homosexuality, and the fall of Oscar Wilde." *AJS; American journal of sociology*, 111(1): 213–248.

Agrawal, Ajay, and Rebecca Henderson. 2002. "Putting patents in context: Exploring knowledge transfer from MIT." *Management science*, 48(1): 44–60.

Alexander, Cindy R. 1999. "On the nature of the reputational penalty for corporate crime: Evidence." *The journal of law & economics*, 42(S1): 489–526.

Allport, Floyd Henry. 1924. "Social psychology (1924)." In *Public Domain Series*. 128–138. Bethlehem, PA:mediastudies.press.

Altonji, Joseph G, Todd E Elder, and Christopher R Taber. 2005. "Selection on observed and unobserved variables: Assessing the effectiveness of catholic schools." *The journal of political economy*, 113(1): 151–184.

Antecol, Heather, and Deborah Cobb-Clark. 2003. "Does Sexual Harassment Training Change Attitudes? A View from the Federal Level." *Social science quarterly*, 84(4): 826–842.

Azoulay, Pierre, Alessandro Bonatti, and Joshua L Krieger. 2017. "The career effects of scandal: Evidence from scientific retractions." *Research policy*, 46(9): 1552–1569.

Azoulay, Pierre, Jeffrey L Furman, Krieger, and Fiona Murray. 2015. "Retractions." *The review of economics and statistics*, 97(5): 1118–1136.

Azoulay, Pierre, Joshua S Graff Zivin, and Jialan Wang. 2010. "Superstar Extinction." *The quarterly journal of economics*, 125(2): 549–589.

Baker, Andrew C, David F Larcker, and Charles C Y Wang. 2022. "How much should we trust staggered difference-in-differences estimates?" *Journal of financial economics*, 144(2): 370–395.

Barmes, Lizzie. 2023. "Silencing at work: Sexual harassment, workplace misconduct and NDAs." *Industrial law journal*, 52(1): 68–106.

Barnett, Michael L, and Andrew A King. 2017. "Good Fences Make Good Neighbors: A Longitudinal Analysis of an Industry Self-Regulatory Institution." *Academy of Management Journal*.

Batut, Cyprien, Caroline Coly, and Sarah Schneider-Strawczynski. 2022. "It's a man's world: culture of abuse, #MeToo and worker flows." *Working Paper*.

Bedera, Nicole Krystine. 2024. *On the wrong side: How universities protect perpetrators and betray survivors of sexual violence.* University of California Press.

Berdahl, Jennifer L. 2007. "Harassment Based on Sex: Protecting Social Status in the Context of Gender Hierarchy." *AMRO*, 32(2): 641–658.

Blackler, Frank. 1995. "Knowledge, knowledge work and organizations: An overview and interpretation." *Organization studies*, 16(6): 1021–1046.

Boudreau, Kevin J, and Nilam Kaushik. 2023. "Gender Differences in Responses to Competitive Organization? A Field Experiment on Differences Between STEM and Non-STEM Fields from an Internet-of-Things Platform." *Organization Science*.

Boudreau, Laura E, Sylvain Chassang, Ada Gonzalez-Torres, and Rachel M Heath. 2023. "Monitoring Harassment in Organizations." *NBER Working Paper*, , (31011).

Bruyaka, Olga, Déborah Philippe, and Xavier Castañer. 2018. "Run away or stick together? The impact of organization-specific adverse events on alliance partner defection." *Academy of management review*, 43(3): 445–469.

Callaway, Brantly, and Pedro H C Sant'Anna. 2021. "Difference-in-Differences with multiple time periods." *Journal of econometrics*, 225(2): 200–230.

Chandra, Amitabh, and Connie Xu. 2025. "Where discovery happens: Research institutions and fundamental knowledge in the life-sciences." National Bureau of Economic Research w33996, Cambridge, MA:National Bureau of Economic Research.

Cheng, Danqiao, Serena Does, Seval Gündemir, and Margaret Shih. 2024. "How organizational responses to sexual harassment claims shape public perception." *Basic and Applied Social Psychology*, 1–18.

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

Chen, Jiafeng, and Jonathan Roth. 2024. "Logs with zeros? Some problems and solutions." *The Quarterly Journal of Economics*, 139(2): 891–936.

Collis, Manuela R, and Clémentine Van Effenterre. 2024. "Workplace Hostility."

Cortina, Lilia M, and Jennifer L Berdahl. 2008. "Sexual harassment in organizations: A decade of research in review." *The SAGE Handbook of Organizational Behavior: Volume I - Micro Approaches*, 469–497.

Cortina, Lilia M, and Maira A Areguin. 2021. "Putting people down and pushing them out: Sexual harassment in the workplace." *Annual review of organizational psychology and organizational behavior*, 8(1): 285–309.

Cortina, Lilia M, and Vicki J Magley. 2003. "Raising voice, risking retaliation: Events following interpersonal mistreatment in the workplace." *Journal of occupational health psychology*, 8(4): 247–265.

Dahl, Gordon B, and Matthew M Knepper. 2021. "Why is Workplace Sexual Harassment Underreported? The Value of Outside Options Amid the Threat of Retaliation." *NBER Working Paper*.

de Chaisemartin, Clément, and Xavier D'Haultfœuille. 2020. "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects." *The American economic review*, 110(9): 2964–2996.

Dionisi, Angela M, and Julian Barling. 2018. "It hurts me too: Examining the relationship between male gender harassment and observers' well-being, attitudes, and behaviors." *Journal of occupational health psychology*, 23(3): 303–319.

Dobbin, Frank, and Alexandra Kalev. 2019. "The promise and peril of sexual harassment programs." *Proceedings of the National Academy of Sciences of the United States of America*, 116(25): 12255–12260.

Dobbin, Frank, and Erin L Kelly. 2007. "How to stop harassment: Professional construction of legal compliance in organizations." *American journal of sociology*, 112(4): 1203–1243.

Durand, Rodolphe, and Jean-Philippe Vergne. 2015. "Asset divestment as a response to media attacks in stigmatized industries: Asset Divestment as a Response to Media Attacks." *Strategic management journal*, 36(8): 1205–1223.

Equal Employment Opportunity Commission. 1980. "Guidelines on Discrimination Because of Sex: Sexual Harassment." 29 C.F.R. § 1604.11(a).

Feldblum, Chai R, and Victoria A Lipnic. 2016. "Select Task Force on the Study of Harassment int he Workplace." U.S. Equal Employment Opportunity Commission.

Fitzgerald, Louise. 2019. "Unseen: the sexual harassment of low-income women in America." *Equality Diversity and Inclusion An International Journal*, 39(1): 5–16.

Fitzgerald, Louise F, and Lilia M Cortina. 2018. "Sexual harassment in work organizations: A view from the 21st century." In *APA handbook of the psychology of women: Perspectives on women's private and public lives, Vol.* Vol. 2, , ed. Cheryl B Travis, 215–234. Washington, DC, US:American Psychological Association, xvi.

Fitzgerald, Louise F, Charles L Hulin, and Fritz Drasgow. 1994. "The antecedents and consequences of sexual harrassment in organizations: An integrated model."

Fitzgerald, Louise F, Sandra L Shullman, Nancy Bailey, Margaret Richards, Janice Swecker, Yael Gold, Mimi Ormerod, and Lauren Weitzman. 1988. "The incidence and dimensions of sexual harassment in academia and the workplace." *Journal of vocational behavior*, 32(2): 152–175.

Fitzgerald, Louise F, Suzanne Swan, and Karla Fischer. 1995. "Why didn't she just report him? The psychological and legal implications of women's responses to sexual harassment." *The Journal of Social Issues*, 51(1): 117–138.

Flecha, Ramón. 2021. "Second-order sexual harassment: Violence against the silence breakers who support the victims." *Violence against women*, 27(11): 1980–1999.

Folke, Olle, and Johanna Rickne. 2022. "Sexual Harassment and Gender Inequality in the Labor Market*." *The quarterly journal of economics*, 137(4): 2163–2212.

Folke, Olle, and Johanna Rickne. 2025. "Sexual harassment across occupations: new evidence from Swedish Nationally representative data." *European sociological review*, jcaf020.

Furman, Jeffrey L, Kyle Jensen, and Fiona Murray. 2012. "Governing knowledge in the scientific community: Exploring the role of retractions in biomedicine." *Research policy*, 41(2): 276–290.

Gertsberg, Marina. 2022. "The unintended consequences of #MeToo - evidence from research collaborations." *SSRN Electronic Journal*.

Goffman, Erving. 1963. *Stigma: notes on the management of spoiled identity. A Spectrum book*, Englewood Cliffs, N.J.:Prentice-Hall.

Gomulya, David, and Warren Boeker. 2016. "Reassessing board member allegiance: CEO replacement following financial misconduct: Reassessing Board Member Allegiance." *Strategic management journal*, 37(9): 1898–1918.

Goodman-Bacon, Andrew. 2021. "Difference-in-differences with variation in treatment timing." *Journal of Econometrics*, 225(2): 254–277.

Graf, Nikki. 2018. "Sexual Harassment at Work in the Era of #MeToo." *Pew Research Center*.

Greve, Henrich R, Donald Palmer, and Jo-Ellen Pozner. 2010. "Organizations gone wild: The causes, processes, and consequences of organizational misconduct." *Academy of Management Annals*, 4(1): 53–107.

Gross, Charles. 2016. "Scientific Misconduct." *Annual review of psychology*, 67: 693–711.

Hershcovis, M Sandy, Babatunde Ogunfowora, Tara C Reich, and Amy M Christie. 2017. "Targeted workplace incivility: The roles of belongingness, embarrassment, and power." *Journal of organizational behavior*, 38(7): 1057–1075.

Hitlan, Robert T, Kimberly T Schneider, and Benjamin M Walsh. 2006. "Upsetting behavior: Reactions to personal and bystander sexual harassment experiences." *Sex roles*, 55(3-4): 187–195.

Ho, Daniel E, Kosuke Imai, Gary King, and Elizabeth A Stuart. 2007. "Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference." *Political analysis: an annual publication of the Methodology Section of the American Political Science Association*, 15(3): 199–236.

Hudson, B. 2008. "Against all Odds: A Consideration of Core-Stigmatized Organizations." *Academy of Management Review*, 33(1): 252–266.

Iacus, Stefano M, Gary King, and Giuseppe Porro. 2011. "Multivariate matching methods that are monotonic imbalance bounding." *Journal of the American Statistical Association*, 106(493): 345–361.

Iacus, Stefano M, Gary King, and Giuseppe Porro. 2012. "Causal inference without balance checking: Coarsened Exact Matching." *Political analysis: an annual publication of the Methodology Section of the American Political Science Association*, 20(1): 1–24.

Ilies, Remus, Nancy Hauserman, Susan Schwochau, and John Stibal. 2003. "Reported incidence rates of work-related sexual harassment in the United States: Using meta-analysis to explain reported rate disparities." *Personnel psychology*, 56(3): 607–631.

Janney, Jay J, and Steve Gove. 2017. "Firm linkages to scandals via directors and professional service firms: Insights from the backdating scandal." *Journal of Business Ethics*, 140(1): 65–79.

Jensen, Michael. 2006. "Should we stay or should we go? Accountability, status anxiety, and client defections." *Administrative Science Quarterly*, 51(1): 97–128.

Johnson, Carrie Ann. 2023. "The purpose of whisper networks: a new lens for studying informal communication channels in organizations." *Frontiers in communication*, 8.

Jones, Edward Ellsworth. 1984. *Social stigma : the psychology of marked relationships.*

Kalev, Alexandra, Frank Dobbin, and Erin Kelly. 2006. "Best practices or best guesses? Assessing the efficacy of corporate affirmative action and diversity policies." *American sociological review*, 71(4): 589–617.

Kang, Eugene. 2008. "Director interlocks and spillover effects of reputational penalties from financial reporting fraud." *Academy of Management journal*, 51(3): 537–555.

Karpoff, Jonathan M, D Scott Lee, and Valaria P Vendrzyk. 1999. "Defense Procurement Fraud, Penalties, and Contractor Influence." *The journal of political economy*, 107(4): 809–842.

Karpoff, Jonathan M, D S Lee, and Gerald S Martin. 2008. "The cost to firms of cooking the books." *Journal of financial and quantitative analysis*, 43(3): 581–611.

Lange, Donald, Peggy M Lee, and Ye Dai. 2011. "Organizational reputation: A review." *Journal of Management*, 37(1): 153–184.

Leskinen, Emily A, and Lilia M Cortina. 2014. "Dimensions of disrespect: Mapping and measuring gender harassment in organizations." *Psychology of women quarterly*, 38(1): 107–123.

Leskinen, Emily A, Lilia M Cortina, and Dana B Kabat. 2011. "Gender harassment: broadening our understanding of sex-based harassment at work." *Law and human behavior*, 35(1): 25–39.

Libarkin, J. 2024. "Academic Sexual Misconduct Database."

Luo, Hong, and Laurina Zhang. 2022. "Scandal, Social Movement, and Change: Evidence from #MeToo in Hollywood." *Management science*, 68(2): 1278–1296.

Lu, Susan Feng, Ginger Zhe Jin, Brian Uzzi, and Benjamin Jones. 2013. "The retraction penalty: evidence from the Web of Science." *Scientific reports*, 3(1): 3146.

Maimone, Giulia, Gil Appel, Craig R M McKenzie, and Ayelet Gneezy. 2025. "Citation penalties following sexual versus scientific misconduct allegations." *PloS one*, 20(3): e0317736.

Maitlis, Sally, and Marlys Christianson. 2014. "Sensemaking in organizations: Taking stock and moving forward." *Academy of Management annals*, 8(1): 57–125.

Maitlis, Sally, and Scott Sonenshein. 2010. "Sensemaking in crisis and change: Inspiration and insights from Weick (1988)." *The journal of management studies*, 47(3): 551–580.

Mansfield, Phyllis Kernoff, Patricia Barthalow Koch, Julie Henderson, Judith R Vicary, Margaret Cohn, and Elaine W Young. 1991. "The job climate for women in traditionally male blue-collar occupations." *Sex roles*, 25(1-2): 63–79.

Manski, C. 2003. *Partial Identification Probability Distributions.* New York:Springer.

McDonnell, Mary-Hunter, and Timothy Werner. 2016. "Blacklisted businesses: Social activists' challenges and the disruption of corporate political activity." *Administrative Science Quarterly*, 61(4): 584–620.

McDonnell, Mary-Hunter, Kate Odziemkowska, and Elizabeth Pontikes. 2021. "Bad Company: Shifts in Social Activists' Tactics and Resources After Industry Crises." *Organization Science*, 32(4): 1033–1055.

McKinley, William, Lawrence A Ponemon, and Allen G Schick. 1996. "Auditors' perceptions of client firms: The stigma of decline and the stigma of growth." *Accounting, organizations and society*, 21(2-3): 193–213.

McLaughlin, Heather, Christopher Uggen, and Amy Blackstone. 2012. "Sexual harassment, workplace authority, and the paradox of power." *American sociological review*, 77(4): 625–647.

McLaughlin, Heather, Christopher Uggen, and Amy Blackstone. 2017. "The Economic and Career Effects of Sexual Harassment on Working Women - Heather McLaughlin, Christopher Uggen, Amy Blackstone, 2017." *Gender & society: official publication of Sociologists for Women in Society.*

Meza, Summer. 2017. "What Is a Whisper Network? How Women Are Taking Down Bad Men in the #MeToo Age." *https://www.newsweek.com/what-whisper-network-sexual-misconduct-allegations-719009*, Accessed: 2025-6-30.

Morral, Andrew R, Kristie L Gore, and Terry L Schell. 2015. *Sexual assault and sexual harassment in the U.s. military: Annex to volume 2. Tabular results from the 2014 rand military workplace study for department of defense service members.* Santa Monica, CA:RAND.

Neu, D, and Michael Wright. 1992. "Bank failures, stigma management and the accounting establishment." *Accounting Organizations and Society*, 17: 645–665.

Oster, Emily. 2019. "Unobservable selection and coefficient stability: Theory and evidence." *Journal of business & economic statistics: a publication of the American Statistical Association*, 37(2): 187–204.

Page, Robert M. 1984. *Stigma.* London, England:Psychology Press.

Piazza, Alessandro, and Julien Jourdan. 2018. "When the dust settles: The consequences of scandals for organizational competition." *Academy of Management journal*, 61(1): 165–190.

Pontikes, Elizabeth, Giacomo Negro, and Hayagreeva Rao. 2010. "Stained red: A study of stigma by association to blacklisted artists during the "red scare" in Hollywood, 1945 to 1960." *American sociological review*, 75(3): 456–478.

Popovich, Paula M, and Michael A Warren. 2010. "The role of power in sexual harassment as a counterproductive behavior in organizations." *Human resource management review*, 20(1): 45–53.

Priem, Jason, Heather Piwowar, and Richard Orr. 2022. "OpenAlex: A fully-open index of scholarly works, authors, venues, institutions, and concepts." *arXiv [cs.DL]*.

Pryor, John B. 1987. "Sexual harassment proclivities in men." *Sex roles*, 17(5-6): 269–290.

Pryor, John B, Christine M LaVite, and Lynnette M Stoller. 1993. "A social psychological analysis of sexual harassment: The person/situation interaction." *Journal of vocational behavior*, 42(1): 68–83.

Raver, Jana L, and Michele J Gelfand. 2005. "Beyond the Individual Victim: Linking Sexual Harassment, Team Processes, and Team Performance." *Academy of Management Journal*.

Rehg, Michael T, Marcia P Miceli, Janet P Near, and James R Van Scotter. 2008. "Antecedents and Outcomes of Retaliation Against Whistleblowers: Gender Differences and Power Relationships." *Organization Science*, 19(2): 221–240.

Rospenda, Kathleen M, Judith A Richman, and Candice A Shannon. 2009. "Prevalence and mental health correlates of harassment and discrimination in the workplace: results from a national study: Results from a national study." *Journal of interpersonal violence*, 24(5): 819–843.

Rowe, Mary P. 1996. "Dealing with Harassment: A Systems Approach." *https://hdl.handle.net/1721.1/155967*, Accessed: 2025-8-15.

Schneider, Kimberly T, Suzanne Swan, and Louise F Fitzgerald. 1997. "Job-related and psychological effects of sexual harassment in the workplace: Empirical evidence from two organizations." *The Journal of Applied Psychology*, 82(3): 401–415.

Shanghai Consultancy. 2024. "ShanghaiRanking Global Ranking of Academic Subjects."

Sims, Carra S, Fritz Drasgow, and Louise F Fitzgerald. 2005. "The effects of sexual harassment on turnover in the military: time-dependent modeling." *The Journal of applied psychology*, 90(6): 1141–1152.

Sockin, Jason, and Aaron Sojourner. 2023. "What's the Inside Scoop? Challenges in the Supply and Demand for Information on Employers." *Journal of Labor Economics*, 41(4): 1041–1079.

Stephan, Paula E. 2010. "The economics of science." In *Handbook of The Economics of Innovation, Vol. 1*. Vol. 1 of *Handbook of the economics of innovation*, 217–273. Elsevier.

Stroube, Bryan K, and Anastasiya Zavyalova. 2024. "The relative effects of a scandal on member engagement in rites of integration and rites of passage: Evidence from a child abuse scandal in the Catholic Archdiocese of Philadelphia." *Organization science*.

Stuart, Toby, and Yanbo Wang. 2016. "Who cooks the books in China, and does it pay? Evidence from private, high-technology firms: Who Cooks the Books in China, and Does It Pay?" *Strategic management journal*, 37(13): 2658–2676.

Subramani, Gauri, and Andreea Gorbatai. 2025. "The Career Costs of Discrimination and Harassment." *Working Paper*.

Sun, Liyang, and Sarah Abraham. 2021. "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects." *Journal of econometrics*, 225(2): 175–199.

Sutton, Robert I, and Anita L Callahan. 1987. "The stigma of bankruptcy: Spoiled organizational image and its management." *Academy of Management journal*, 30(3): 405–436.

Trawalter, Sophie, Jennifer Doleac, Lindsay Palmer, Kelly Hoffman, and Adrienne Carter-Sowell. 2022. "Women's Safety Concerns and Academia: How Safety Concerns Can Create Opportunity Gaps." *Social psychological and personality science*, 13(2): 403–415.

Wang, Yanbo, Toby Stuart, and Jizhen Li. 2021. "Fraud and Innovation." *Administrative science quarterly*, 66(2): 267–297.

Weick, Karl E. 1988. "ENACTED SENSEMAKING IN CRISIS SITUATIONS[1]." *The journal of management studies*, 25(4): 305–317.

Weick, Karl E. 1995. *Sensemaking in Organizations. Foundations for Organizational Science*, Thousand Oaks, CA:SAGE Publications.

Weick, Karl E, Kathleen M Sutcliffe, and David Obstfeld. 2005. "Organizing and the process of sensemaking." *Organization science*, 16(4): 409–421.

Welsh, Sandy. 1999. "Gender and Sexual Harassment." *Annual review of sociology*, 25: 169–190.

Widmann, Rainer, Michael Rose, and Marina Chugunova. 2022. "Allegations of sexual misconduct, accused scientists, and their research." *SSRN Electronic Journal*.

Wiesenfeld, B, Kurt Wurthmann, and D Hambrick. 2008. "The stigmatization and devaluation of elites associated with corporate failures: A process model." *Academy of Management Review*, 33(1): 231–251.

Willness, Chelsea R, Piers Steel, and Kibeom Lee. 2007. "A meta-analysis of the antecedents and consequences of workplace sexual harassment." *Personnel psychology*, 60(1): 127–162.

Yenkey, Chris, and Donald Palmer. 2025. "Consequences of performance-enhancing misconduct: Insights from professional road cycling, 2000–2010." *Management science*.

Yenkey, Christopher B. 2018. "Fraud and Market Participation: Social Relations as a Moderator of Organizational Misconduct." *Administrative science quarterly*, 63(1): 43–84.

Zhang, Lin, Gunnar Sivertsen, Huiying Du, Ying Huang, and Wolfgang Glänzel. 2021. "Gender differences in the aims and impacts of research." *Scientometrics*, 126(11): 8861–8886.

(For Online Publication)

Appendix to

Sexual Misconduct and Scientific Production

Manuela R. Collis

List of Appendices

Appendix A: Clarivate Core Collection **57**

Appendix B: Tables **59**

Appendix C: Additional Results **64**

Appendix D: Robustness of the Null Finding - Full Outline **66**

A Clarivate Core Collection

I use nine primary departments which are based on Clarivate's Core Collection (`https://mjl.clarivate.com/help-center`): Agriculture, Biology & Environmental Sciences; Arts & Humanities; Business Collection; Clinical Medicine; Electronics & Telecommunication Collection; Engineering, Computing & Technology; Life Sciences; Physical, Chemical & Earth Sciences; Social And Behavioral Sciences. Below, I list for the research communities defined by Clarivate and SJR for each of the nine disciplines.

Disciplines and Communities

Agriculture, Biology & Environmental Sciences
- Agriculture/Agronomy
- Agricultural Chemistry
- Animal Sciences
- Aquatic Sciences
- Biology
- Biodiversity
- Biophysics
- Biotechnology
- Botany
- Conservation
- Developmental Biology
- Ecology/Environmental Sciences
- Entomology
- Evolutionary Biology
- Fisheries
- Food Science
- Forestry
- Horticulture
- Marine Biology
- Molecular Biology
- Mycology
- Paleontology
- Parasitology
- Plant Sciences
- Soil Science
- Veterinary Sciences
- Wildlife Management
- Zoology

Arts & Humanities
- Architecture
- Art
- Asian Studies
- Classical Studies
- Dance
- Film, Radio & TV
- Folklore
- History
- Humanities (General)
- Language & Linguistics
- Literary Criticism
- Literature
- Medieval Studies
- Music
- Philosophy
- Poetry
- Religion
- Renaissance Studies
- Theater

Business Collection
- Accounting
- Advertising
- Banking
- Business
- E-commerce
- Economics
- Finance
- Hospitality Industry
- Human Resources
- Insurance
- International Business
- Logistics
- Management
- Marketing
- Nonprofit Organizations
- Operations Research
- Real Estate

Clinical Medicine
- Allergy
- Anesthesiology
- Audiology
- Cardiology
- Critical Care
- Dentistry
- Dermatology
- Emergency Medicine
- Endocrinology
- Gastroenterology
- General Medicine
- Geriatrics
- Health Policy
- Hematology
- Immunology
- Infectious Diseases
- Medical Ethics
- Medical Informatics
- Medical Technology
- Neurology
- Nursing
- Nutrition
- Obstetrics & Gynecology
- Oncology
- Ophthalmology
- Orthopedics
- Pathology
- Pediatrics
- Pharmacology
- Physical Therapy
- Preventive Medicine
- Psychiatry
- Radiology
- Rheumatology
- Sports Medicine
- Surgery
- Toxicology
- Transplantation
- Tropical Medicine
- Urology

Electronics & Telecommunications Collection
- Artificial Intelligence
- Automation
- Computer Hardware
- Computer Science

- Control Systems
- Cybernetics
- Digital Signal Processing
- Electrical Engineering
- Electronics
- Embedded Systems
- Information Systems
- Machine Learning
- Mobile Communications
- Network Security
- Robotics
- Software Engineering
- Telecommunications

Engineering, Computing & Technology

- Acoustics
- Aerospace Engineering
- Architectural Engineering
- Automotive Engineering
- Biomedical Engineering
- Ceramics
- Chemical Engineering
- Civil Engineering
- Composite Materials
- Construction
- Energy & Fuels
- Environmental Engineering
- Fluid Dynamics
- Industrial Engineering
- Manufacturing
- Marine Engineering
- Materials Science
- Mechanical Engineering
- Metallurgy
- Mining Engineering
- Nanotechnology
- Nuclear Engineering
- Petroleum Engineering
- Polymers
- Remote Sensing
- Thermodynamics
- Transportation
- Water Resources

Life Sciences

- Biochemistry
- Biomedical Research
- Biometrics
- Cancer Research
- Cell Biology
- Computational Biology
- Genetics
- Genomics
- Immunobiology
- Limnology
- Microbiology
- Microscopy
- Molecular Biology
- Neuroscience
- Physiology
- Proteomics
- Stem Cell Research
- Structural Biology
- Systems Biology
- Tissue Engineering
- Toxinology
- Virology

Physical, Chemical & Earth Sciences

- Analytical Chemistry
- Applied Mathematics
- Astronomy & Astrophysics
- Atmospheric Science
- Atomic Physics
- Catalysis
- Chemistry
- Computational Physics
- Condensed Matter Physics
- Crystallography
- Electrochemistry
- Fluid Dynamics
- Geochemistry
- Geology
- Geophysics
- Inorganic Chemistry
- Mathematics
- Meteorology
- Mineralogy
- Nanotechnology
- Nuclear Physics
- Oceanography
- Optics
- Organic Chemistry
- Particle Physics
- Physical Chemistry
- Polymer Science
- Quantum Physics
- Seismology
- Spectroscopy
- Statistical Physics
- Thermodynamics

Social And Behavioral Sciences

- Anthropology
- Applied Linguistics
- Archaeology
- Area Studies
- Behavioral Sciences
- Child Development
- Clinical Psychology
- Cognitive Science
- Communication
- Criminology
- Cultural Studies
- Demography
- Developmental Psychology
- Economic Geography
- Education
- Educational Psychology
- Environmental Studies
- Ergonomics
- Ethics
- Ethnic Studies
- Experimental Psychology
- Family Studies
- Geography
- Gerontology
- Health Education
- Human Geography
- Industrial Psychology
- Information Science
- International Relations
- Law
- Library Science
- Political Science
- Psychoanalysis
- Psychology
- Public Administration
- Social Psychology
- Social Work
- Sociology
- Special Education
- Sports Science
- Substance Abuse
- Urban Studies
- Women's Studies

B Tables

Table B1: Outcome Descriptive Statistics Before and After Matching

	Before Matching		**After Matching**	
Variable	**Control**	**Treated**	**Control**	**Treated**
Publications per Capita Total				
Mean (SD)	2.18 (2.65)	1.53 (1.11)	2.98 (3.53)	1.51 (1.02)
N	50,188	210	4,759	202
Publications per Capita Women				
Mean (SD)	2.28 (2.73)	1.51 (1.11)	2.99 (3.47)	1.48 (0.94)
N	33,632	188	3,492	181
Publications per Capita Men				
Mean (SD)	2.33 (2.94)	1.64 (1.42)	3.14 (3.76)	1.59 (1.28)
N	44,991	198	4,268	190

Note: Control group statistics after matching use analytic weights from CEM.

Table B2: L1 Imbalance Before and After Matching at time period $t-1$

Variable	**Before Matching**	**After Matching**	**Reduction (%)**
Multivariate L1 distance	0.710	0.501	29.44
Univariate L1 distance			
Discipline	0.293	0.135	53.92
Public Institution	0.112	2.2e-15	100.00
State Law for Hostility & Harassment	0.081	4.1e-15	100.00
Shanghai Ranking (GRAS)	0.032	5.3e-16	100.00
Department Size	0.445	0.147	66.93

Note: The variables *Discipline*, *Public Institution* and *State Laws for Discrimination & Harassment* are categorical.

Table B3: Effect of Occurrence of Sexual Misconduct on Average Number of Publications per Department Member

	Average number of publications per department member				
	(1) Full Sample	(2) 5 yr delay	(3) Adj. control	(4) Women	(5) Men
5 Year Average TE	0.022	0.013	-0.019	0.018†	-0.032
	(0.023)	(0.020)	(0.029)	(0.024)	(0.023)
6 periods before treatment	-0.017	0.015	-0.019	-0.024	-0.011
	(0.021)	(0.037)	(0.030)	(0.037)	(0.046)
5 periods before treatment	0.019	-0.013	-0.008	0.031	0.028
	(0.025)	(0.038)	(0.033)	(0.049)	(0.048)
4 periods before treatment	-0.022	-0.012	-0.030	-0.035	-0.029
	(0.021)	(0.038)	(0.028)	(0.030)	(0.044)
3 periods before treatment	0.022	0.066	-0.011	-0.009	0.018
	(0.026)	(0.042)	(0.032)	(0.033)	(0.043)
2 periods before treatment	0.051†	0.083†	0.026	0.033	0.023
	(0.029)	(0.044)	(0.043)	(0.029)	(0.039)
1 periods before treatment	(.)	(.)	(.)	(.)	(.)
Event period (t=0)	0.007	0.001	0.004	0.009	-0.037
	(0.026)	(0.036)	(0.038)	(0.032)	(0.033)
1 periods after treatment	0.002	0.001	-0.004	0.010	-0.013
	(0.025)	(0.033)	(0.036)	(0.037)	(0.033)
2 periods after treatment	0.079	0.003	-0.013	0.093	-0.012
	(0.067)	(0.041)	(0.039)	(0.077)	(0.037)
3 periods after treatment	0.002	0.035	-0.031	-0.015	-0.053†
	(0.027)	(0.040)	(0.040)	(0.031)	(0.031)
4 periods after treatment	0.020	0.032	-0.046	-0.008	-0.067
	(0.031)	(0.035)	(0.038)	(0.030)	(0.042)
5 periods after treatment	0.013	-0.002	-0.021	-0.012	-0.061
	(0.022)	(0.030)	(0.036)	(0.033)	(0.039)
Observations	53,841	35,544	12,691	29,623	35,880
R-squared	0.9385	0.9360	0.9767	0.9465	0.8907

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Standard errors clustered at university x department level. Fixed effects for university x department and years. Column 2 restricts the sample to incidents for which public reporting occurs after more than five years. Column 3 restricts the control group to departments which have a representation of women that is above the median for a given discipline.

Table B4: Comparison of Departments with Varying Degrees of Ability to Switch

	Occurrence of Sexual Misconduct		
	(1) Full Sample	(2) Hard to switch	(3) Professional / Clinical medicine
5 Year Average TE	0.022	0.037	-0.12**
	(0.023)	(0.025)	(0.034)
6 periods before treatment	-0.017	-0.018	-0.012
	(0.021)	(0.023)	(0.055)
5 periods before treatment	0.019	0.023	-0.032
	(0.025)	(0.027)	(0.053)
4 periods before treatment	-0.022	-0.017	-0.079*
	(0.021)	(0.022)	(0.034)
3 periods before treatment	0.022	0.024	-0.017
	(0.026)	(0.028)	(0.051)
2 periods before treatment	0.051†	0.060†	-0.040
	(0.029)	(0.032)	(0.040)
1 periods before treatment	(.)	(.)	(.)
Event period (t=0)	0.007	0.015	-0.067
	(0.026)	(0.028)	(0.044)
1 periods after treatment	0.002	0.011	-0.091*
	(0.025)	(0.028)	(0.040)
2 periods after treatment	0.079	0.099	-0.113*
	(0.067)	(0.075)	(0.052)
3 periods after treatment	0.002	0.018	-0.143*
	(0.027)	(0.028)	(0.063)
4 periods after treatment	0.020	0.036	-0.131*
	(0.031)	(0.034)	(0.059)
5 periods after treatment	0.013	0.031	-0.178**
	(0.022)	(0.024)	(0.062)
Observations	53,841	42,122	11,719
R-squared	0.9385	0.9397	0.9327

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Standard errors clustered at university x department level. Fixed effects for university x department and years.

Table B5: Effect of Public Reporting on Number of Publications per Department Member

	Simple DiD (1)	Simple Event-study (2)	Matched DiD (3)	Matched Event-study (4)	Sun & Abraham (5)	S&A 5 yr delay (6)
5 Year Average TE	-0.0402*	-0.0443**	-0.0950†	-0.1079***	-0.0972***	-0.1900***
	(0.0186)	(0.0139)	(0.0521)	(0.0289)	(0.0272)	(0.0485)
5 years before treatment		0.0166		-0.0117	-0.0154	-0.0821†
		(0.0207)		(0.0259)	(0.0276)	(0.0435)
4 years before treatment		0.0231		-0.0075	-0.0086	-0.0606
		(0.0227)		(0.0287)	(0.0290)	(0.0502)
3 years before treatment		0.0084		-0.0187	-0.0134	-0.0628
		(0.0223)		(0.0265)	(0.0263)	(0.0393)
2 years before treatment		-0.0104		-0.0460	-0.0598*	-0.1095*
		(0.0247)		(0.0304)	(0.0287)	(0.0484)
1 years before treatment		(.)		(.)	(.)	(.)
Event period (t=0)		-0.0107		-0.0502	-0.0386	-0.1315†
		(0.0245)		(0.0397)	(0.0339)	(0.0714)
1 year after treatment		-0.0394		-0.0983†	-0.0905†	-0.1772*
		(0.0250)		(0.0519)	(0.0477)	(0.0856)
2 years after treatment		-0.0523*		-0.1050†	-0.1142*	-0.1784*
		(0.0232)		(0.0544)	(0.0509)	(0.0755)
3 years after treatment		-0.0742**		-0.1440*	-0.1327*	-0.2319**
		(0.0247)		(0.0656)	(0.0612)	(0.0884)
4 years after treatment		-0.0421		-0.1156	-0.0969	-0.2512†
		(0.0390)		(0.0802)	(0.0764)	(0.1301)
5 years after treatment		-0.0137		-0.0766	-0.0517	-0.1113
		(0.0390)		(0.0671)	(0.0638)	(0.1453)
Dept. Member Count (lagged)	0.0008*	0.0008*	0.0003	0.0003		
	(0.0003)	(0.0003)	(0.0004)	(0.0004)		
Observations	53,481	53,481	39,381	39,381	39,484	33,314
Clusters	2,086	2,086	1,485	1,485	1,494	1,304
R-squared	0.938	0.938	0.936	0.936	0.936	0.936
Unit FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Cohort-Specific FE	No	No	No	No	No	No

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Dependent variable is publications per capita at the department level. Standard errors clustered at the department level in parentheses. Never-treated departments used as control group. Columns (3)-(5) use CEM-matched sample with analytical weights. Event study columns show dynamic treatment effects. Sun & Abraham column uses interaction-weighted estimator. The "5 Year Average TE" row shows the average treatment effect over years 1-5 post-treatment for event study columns. All specifications include department and year fixed effects.

Table B6: The Effect of Public Reporting on Publication Output by Gender

	Public Reporting of Sexual misconduct		
	(1) Total	(2) Women	(3) Men
5 Year Average TE	-0.095*	-0.06**	-0.11†
	(0.048)	(0.016)	(0.025)
6 periods before treatment	-0.035	-0.037	-0.041
	(0.032)	(0.032)	(0.044)
5 periods before treatment	-0.017	-0.048	-0.030
	(0.026)	(0.034)	(0.035)
4 periods before treatment	-0.011	-0.050	0.002
	(0.032)	(0.026)	(0.49)
3 periods before treatment	-0.011	-0.017	-0.013
	(0.027)	(0.034)	(0.032)
2 periods before treatment	-0.063	-0.012	-0.091**
	(0.029)	(0.036)	(0.033)
1 periods before treatment	0.000	0.000	0.000
	(.)	(.)	(.)
Event period (t=0)	-0.043	-0.018	-0.044
	(0.034)	(0.033)	(0.04)
1 periods after treatment	-0.091†	-0.048	-0.087
	(0.047)	(0.037)	(0.054)
2 periods after treatment	-0.113*	-0.096***	-0.125*
	(0.049)	(0.031)	(0.062)
3 periods after treatment	-0.132*	-0.099***	-0.157†
	(0.061)	(0.029)	(0.085)
4 periods after treatment	-0.094	-0.059***	-0.159
	(0.075)	(0.046)	(0.104)
Observations	39,484	29,623	38,240
R-squared	0.9361	0.9448	0.8311

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Standard errors clustered at university x department level. Fixed effects for university x department and years.

C Tables

Table C7: Effect of Sexual Misconduct on Publication per Department Member - Delay in Public Reporting

	(1) Same-year	(2) within 0-5 yrs	(3) after 6-10 yrs
5 Year Average TE	-0.058	0.016	0.012
	(0.070)	(0.033)	(0.043)
6 periods before treatment	-0.100	-0.012	-0.056
	(0.084)	(0.031)	(0.056)
5 periods before treatment	-0.020	0.079†	-0.060
	(0.084)	(0.041)	(0.068)
4 periods before treatment	-0.010	0.007	-0.093
	(0.108)	(0.028)	(0.069)
3 periods before treatment	0.039	0.026	0.079
	(0.092)	(0.038)	(0.073)
2 periods before treatment	-0.136†	0.042	0.026
	(0.076)	(0.042)	(0.068)
1 periods before treatment	(.)	(.)	(.)
Event period (t=0)	-0.008	0.021	-0.049
	(0.100)	(0.038)	(0.051)
1 periods after treatment	-0.112	-0.000	0.001
	(0.115)	(0.042)	(0.058)
2 periods after treatment	-0.196†	0.123	0.019
	(0.107)	(0.123)	(0.066)
3 periods after treatment	-0.176	-0.052	0.055
	(0.119)	(0.038)	(0.070)
4 periods after treatment	0.205	-0.009	0.036
	(0.266)	(0.051)	(0.063)
5 periods after treatment	0.032	0.009	-0.036
	(0.085)	(0.037)	(0.036)
Observations	668	4,091	1,317
R-squared	0.9747	0.8978	0.8896

$^{\dagger}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. Standard errors clustered at university x department level. Fixed effects for university x department and years. Column 3 (Poisson model) reports Pseudo R-squared in place of R-squared.

D Robustness of the Null Finding - Full Outline

As reported in Figure 4, the occurrence of sexual misconduct shows no discernible effect on the productivity of university departments. This is a lower bound effect since the sampling process is such that only cases that eventually become public are in my sample of treated departments. This means I have a contaminated control group and this likely biases my estimates downwards. In other words, the treated group consists of departments where sexual misconduct occurred (identified ex post via public reporting). The control group includes departments with no known misconduct and departments with misconduct that never became public. Therefore, contamination in the control group would attenuate any true negative effects toward zero.

This leaves room for the possibility, in the extreme case, that we compare departments where sexual misconduct occurred and became public, with departments where sexual misconduct occurred but never became public. Intuitively, we would compare two very similar groups of departments with each other—at time of occurrence—and thus, the null effect may seem sensible, but potentially not estimating the treatment effect of interest, which is the effect of the occurrence of sexual misconduct in university departments on department-level productivity.

Theoretically, we would expect that the occurrence of sexual misconduct in university departments has either a negative effect or a null effect on the productivity of departments. On one hand, there are reasons to believe sexual misconduct could negatively affect organizational productivity through several mechanisms: victims share experiences through whisper networks which can affect task conflict and performance, bystanders may experience psychological consequences that impact their work, and organizational responses like investigations and remedial measures (training programs, new policies, audits) can be disruptive and divert resources from core functions. On the other hand, there are reasons to expect no discernible effects due to significant barriers to information flow within organizations: many incidents go unreported due to fear of stigma and retaliation, organizational procedures emphasize confidentiality and often include nondisclosure agreements, and even when some members learn of incidents through informal channels, pluralistic ignorance and fear of scandal may prevent open discussion or coordinated action, ultimately containing the impact to those directly

involved rather than affecting broader organizational performance.

Therefore, the interpretation of the estimated effect depends on two factors: the contamination level in the control group and the true effect size. As a starting point, we may want to ask how high contamination has to be in order for the results to be non-negligible. Learning more about the level of contamination required, will give us a better sense of whether the observed negligible effect is sensible or not. What we will see, is that the required levels of contamination are far higher than what empirical evidence would suggest.

While the inquiry of the required level of contamination provides us with useful insights, we can also investigate the following four scenarios related to the contamination levels and the true effect size. This allows us to evaluate how possible it is that there is a true null effect: (A) low contamination with negligible true effects, (B) low contamination with meaningfully negative true effects, (C) high contamination with negligible true effects, and (D) high contamination with meaningfully negative true effects. Below, I formally test which scenarios are consistent with the observed data, using a contamination adjustment model to evaluate whether the null finding reflects a genuine absence of productivity impacts or merely contamination-induced attenuation. This approach is in spirit of Manski (2003), Altonji, Elder and Taber (2005), and Oster (2019), who consider how selection into control or treatment affects stability of the observed treatment effect.

Conceptual Framework

The interpretation of the null finding for occurrence effects depends critically on the level of contamination in the control group—that is, the proportion of control departments that experienced unreported sexual misconduct. To establish plausible bounds for this parameter, I need to carefully distinguish between two related but distinct concepts: the total prevalence of misconduct across all departments, and the contamination rate within the control group specifically.

Boudreau et al. (2023) document that 38% of teams in their study experienced at least one case of sexual harassment. This represents the *total prevalence*—the proportion of all departments experiencing misconduct, whether reported or not. In my sample, this total prevalence includes both the 261 treated departments (where misconduct became public) and any control departments with unreported misconduct. To translate between total prevalence

and contamination rates, I use the relationship: Total Prevalence $= P(\text{Treated}) + P(\text{Control}) \times c$, where $P(\text{Treated}) = 261/2,101 = 0.124$ and $P(\text{Control}) = 1,840/2,101 = 0.876$.

If the total prevalence in my setting matches Boudreau et al. (2023)'s 38%, this implies a contamination rate of $c = (0.38 - 0.124)/0.876 = 0.29$, or approximately 29% of control departments having unreported misconduct. For this analysis, I define "low contamination" as $c = 0.10$, representing a scenario where institutional reporting mechanisms function effectively and most serious misconduct either becomes public or does not occur. In contrast, "high contamination" is set at $c = 0.29$, matching the level implied by empirical evidence. To put this in perspective, a contamination rates of 14% in our context would mean unreported cases roughly equal reported ones, while the 29% contamination implied by empirical evidence suggests unreported cases outnumber reported ones by two-to-one across all research universities.

To distinguish economically meaningful from negligible effects, I pre-specify a smallest effect size of interest (SESOI) of 0.027 publications per department and year (≈0.49 papers per year in a median-sized department of 18 members). I classify effects at or below -0.056 as meaningfully negative for departmental productivity (one full publication per year in an 18-faculty department). With 261 treated and 1,840 control departments and a standard error of 0.023, the minimum detectable effect (MDE) at 80% power is 0.064 publications per faculty-year.

The contamination adjustment model makes several key assumptions: (1) contamination affects only control departments, not treated ones; (2) the occurrence effect is homogeneous across departments; and (3) there are no other sources of bias such as baseline imbalances, differential time trends, or measurement lags. Under these assumptions, contamination mechanically attenuates the observed treatment-control contrast. When a fraction c of the control departments experienced unreported misconduct, the observed difference is diluted relative to the true effect. To see why, consider what we're actually comparing:

Treated departments: All experienced a case of sexual misconduct. If the true effect of misconduct is τ (average change in publications per department over five years after treatment), these departments have an average outcome of $Y_0 + \tau$, where Y_0 is the baseline productivity.

Control departments: A mixture of:

- Fraction $(1 - c)$: Truly unaffected departments with outcome Y_0

- Fraction c: Contaminated departments (unreported misconduct) with outcome $Y_0 + \tau$

The control group's average outcome is $(1-c) \times Y_0 + c \times (Y_0 + \tau) = Y_0 + c\tau$, where fraction $(1-c)$ are truly unaffected departments with outcome Y_0 and fraction c are contaminated departments with outcome $Y_0 + \tau$. The observed treatment-control difference is therefore $(Y_0 + \tau) - (Y_0 + c\tau) = \tau - c\tau = \tau(1-c)$. This yields the contamination adjustment formula:

$$\text{True Effect} = \frac{\text{Observed Effect}}{1-c} \tag{6}$$

This formula shows that as contamination increases, the observed effect shrinks toward zero, potentially masking real impacts. A concrete illustration clarifies the logic under empirically-grounded contamination with a meaningful negative effect. Consider 100 control departments where 29% have unreported misconduct ($c = 0.29$) and the true effect is -0.056 publications per faculty member. The 71 departments without misconduct experience no change (0), while the 29 departments with unreported misconduct experience the full effect (-0.056). The average change in the control group becomes $(71 \times 0 + 29 \times (-0.056))/100 = -0.016$. Comparing treated departments (all at -0.056) to this mixed control group (averaging -0.016) yields an observed difference of $-0.056 - (-0.016) = -0.040$, not the true -0.056. The effect is compressed by 29%. Under this contamination-only model, negative true effects would produce negative expected observed effects (unless contamination exceeds 100%, which is outside of possible levels of contamination).

What level of contamination will generate a non-negligible effect?

Given the observed effect of 0.022, what levels of contamination would we need to assume for there to be a true significant negative effect (-0.056)? We can use our contamination adjustment formula, True Effect $\approx$ Observed Effect / $(1-c)$, and rearrange it to solve for c. Starting with $-0.056 \approx 0.022/(1-c)$, we multiply both sides by $(1-c)$ to get $-0.056 \times (1-c) \approx 0.022$. Expanding this gives us $-0.056 + 0.056c \approx 0.022$. Adding 0.056 to both sides yields $0.056c \approx 0.078$. Finally, dividing both sides by 0.056 gives us $c \approx 0.078/0.056 \approx 1.393$.

A value of $c > 1$ suggests that the contamination would have to exceed 100% for there to be possibly a significant negative effect. Before drawing any conclusions, recall that our estimate is not a precisely estimated zero. Instead, the observed effect of 0.022 is somewhat noisy. We

can take this into account and compute a lower-bound and upper-bound contamination level.

The calculation of contamination bounds incorporates the uncertainty in the estimate. We will use it to determine the range of contamination levels that could potentially reconcile the observed estimate with a meaningful negative true effect. That is, instead of simply relying on the estimate, we can take advantage of the known standard errors (0.023) to compute the confidence interval and use the full range of plausible values. The 95% confidence interval for the observed effect spans from -0.023 to 0.067. We can use this information to calculate a contamination bound. We can assume the observed effect is -0.023 and again compute the contamination levels required for such an effect size. And we can do the same, assuming the observed effect size is the upper level of the confidence interval, 0.067.

At the lower bound of -0.023, the required contamination would be $c = 1 - (-0.023/-0.056) = 0.589$, or about 59%. This means that even if the true observed effect were at the very bottom of what's statistically plausible, we would still need nearly 60% of the control departments to have experienced unreported misconduct for the data to be consistent with meaningful productivity losses. At the upper bound of 0.067, the calculation yields $c = 2.196$. Since a contamination level of above 1 is outside of the possible levels of contamination, this scenario can be ruled out. Therefore, we can say that for there to plausibly be a meaningful negative effect, the contamination levels will have to be at least 59%.

To evaluate how plausible this required contamination level is, we can translate it to total prevalence and compare with empirical evidence. If 59% of our 1,840 control departments have unreported misconduct, this represents 1,086 departments. Adding our 261 treated departments yields 1,347 total departments with misconduct out of 2,101 departments—a total prevalence of 64%. This far exceeds the 38% prevalence documented by Boudreau et al. (2023). Conversely, if we accept their 38% prevalence as accurate for our setting, the implied contamination in our control group would be only $(0.38 - 0.124)/0.876 = 29\%$—well below the 59% threshold needed for meaningful negative effects.

This analysis offers us two insights. First, even though the estimated effect is somewhat noisy, the contamination bounds allow us to take that into account and provide us with a useful bound of contamination level c under which the true effect size would be meaningfully negatively affecting productivity. That contamination bound ranges from 59% to 100%.

Second, this range implies that the total prevalence of sexual misconduct would need to

be at least 64%. However, work by Boudreau et al. (2023) suggests that the true prevalence is closer to 38%, which would correspond to only 29% contamination in our control group. This suggests that the observed negligible effect may indeed be a true negligible effect.

Testing whether the true effect could be negligible (Scenarios A & C)

Given the observed effect of 0.022, what would the true effect be under different contamination levels? Contamination attenuates the observed effect according to the formula: Observed Effect = True Effect $\times$ $(1-c)$. Rearranging this gives us True Effect = Observed Effect / $(1-c)$, which allows us to recover the true effect from our observed estimate.

Scenario A: Low contamination ($c = 0.10$)

If only 10% of control departments have unreported misconduct, we can calculate the adjusted effect:

$$\text{True Effect} = \frac{0.022}{1-0.10} = \frac{0.022}{0.90} = 0.024 \tag{7}$$

This adjusted effect of 0.024 publications per faculty-year falls below our pre-specified SESOI threshold of 0.027. In a median department of 18 faculty, this translates to less than half a publication per year (0.43 papers)—a change that would be difficult to detect in practice and unlikely to meaningfully affect departmental operations.

To account for statistical uncertainty, we can also compute the 95% confidence interval for the adjusted effect. Our original confidence interval spans from -0.023 to 0.067. Since we treat c as known and not as a random variable, we are able to adjust the confidence interval of the initial result linearly. That is, we can divide the lower bound and upper bound of the confidence interval by $1-c$ to obtain the confidence interval for the estimated effect under the scenario where the contamination level is 10%. We compute $-0.023/0.90$ and $0.067/0.90$ which gives us a 95% confidence interval of $[-0.026, 0.074]$ for this scenario.

Scenario C: High contamination ($c = 0.29$)

If 29% of control departments have unreported misconduct—matching the contamination level implied by Boudreau et al. (2023)'s documented prevalence—the adjustment becomes:

$$\text{True Effect} = \frac{0.022}{1-0.29} = \frac{0.022}{0.71} = 0.031 \tag{8}$$

This adjusted effect of 0.031 publications per faculty-year marginally exceeds our pre-

specified SESOI threshold of 0.027. In practical terms, this represents 0.56 publications per year in a median 18-faculty department—approximately one paper every two years. While this effect exceeds the SESOI threshold, it remains modest in magnitude and may not be economically significant for departmental decision-making or resource allocation.

To account for statistical uncertainty, we compute the 95% confidence interval for the adjusted effect. Our original confidence interval spans from -0.023 to 0.067. Since we treat c as known and not as a random variable, we can adjust the confidence interval of the initial result linearly by dividing the lower and upper bounds by $1 - c$. We compute $-0.023/0.71$ and $0.067/0.71$, which gives us a 95% confidence interval of $[-0.032, 0.094]$ for this scenario. Notably, this confidence interval includes zero, suggesting the adjusted effect is not statistically distinguishable from null.

Testing whether meaningful negative effects are plausible (Scenarios B & D)

If the true effect were meaningfully negative ($\tau = -0.056$), what would we expect to observe? We can work backwards from this assumption to calculate what we should have seen in our data, then test whether our actual observation is consistent with these predictions.

To test whether our observed value is statistically consistent with each scenario's prediction, I need to determine how likely it is to observe our estimate (0.022) if the scenario were true. Since our observed effect is a sample estimate with standard error 0.023, it follows approximately a normal distribution. Under the null hypothesis that a given scenario is correct (e.g., the expected value is -0.050), the difference between our observed and expected values, standardized by the standard error, follows a standard normal distribution: $z = \frac{\text{Observed} - \text{Expected}}{SE}$. This z-statistic measures how many standard errors our observation lies from the scenario's prediction. Large absolute values of z (typically $|z| > 1.96$) indicate that our observation would be unlikely if the scenario were true, allowing us to assess each scenario's plausibility.

Scenario B: Low contamination ($c = 0.10$)

If the true effect is -0.056 and only 10% of control departments have unreported misconduct, we can calculate what we would expect to observe. Recall that contamination attenuates the observed effect by a factor of $(1 - c)$, so our expected observation would be $-0.056 \times (1 - 0.10) = -0.056 \times 0.90 = -0.050$. This means we should observe departments

with public misconduct having 0.050 fewer publications per faculty-year compared to the control group. Our actual observation is 0.022. To test whether this discrepancy could arise by chance, we calculate:

$$z = \frac{0.022 - (-0.050)}{0.023} = \frac{0.072}{0.023} = 3.13 \tag{9}$$

which yields $p < 0.01$. This indicates that our observation is highly unlikely if this scenario were true.

Scenario D: High contamination ($c = 0.29$)

If the true effect is -0.056 and 29% of control departments have unreported misconduct, the expected observation becomes $-0.056 \times (1 - 0.29) = -0.056 \times 0.71 = -0.040$. Even with empirically-grounded contamination diluting the effect, we should still observe treated departments being less productive by 0.040 publications per faculty-year. Our actual observation is 0.022. To test whether this discrepancy could arise by chance, we calculate:

$$z = \frac{0.022 - (-0.040)}{0.023} = \frac{0.062}{0.023} = 2.70 \tag{10}$$

which yields $p < 0.01$, indicating that our observation is statistically inconsistent with this scenario.

The analysis addresses two key questions that together determine the interpretation of our null finding. First, can the true effect be negligible given what we observed? Scenarios A and C examined this question. After adjusting for contamination, the true effects would be 0.024 (low contamination) or 0.031 (high contamination). In Scenario A, the adjusted effect falls below the SESOI threshold of 0.027. In Scenario C, while the point estimate marginally exceeds SESOI, the confidence interval $[-0.032, 0.094]$ includes zero. Both scenarios remain consistent with negligible to modest true effects.

Second, are meaningful negative effects plausible given what we observed? Scenarios B and D examined this question. If misconduct truly reduced productivity by 0.056 publications per faculty-year, we would expect to observe values of -0.050 under low contamination or -0.040 under high contamination. Our observed estimate of 0.022, while somewhat noisy (SE = 0.023), is 3.13 and 2.70 standard errors away from these predictions, respectively. These statistical tests suggest that meaningful negative effects are unlikely given our data ($p < 0.01$ and $p < 0.01$).

Summary

This multi-pronged analysis offers evidence that the observed effect is a true null effect and that sexual misconduct has, indeed, negligible effects on departmental productivity. Three complementary approaches converge on this conclusion. First, the contamination bounds analysis shows that even accounting for the noisiness in our data, meaningful negative effects would require at least 59% contamination in control departments—corresponding to 64% total prevalence of misconduct across all departments, far exceeding the 38% documented by Boudreau et al. (2023). Moreover, if we accept their 38% prevalence as accurate, this implies only 29% contamination in our control group, well below the required threshold. Second, when we adjust for plausible contamination levels (10–29%), the true effects remain modest, ranging from 0.024 to 0.031 publications per faculty-year, with confidence intervals that include zero and exclude meaningful productivity losses. Third, statistical tests reject scenarios where true effects are meaningfully negative: if misconduct truly reduced productivity by 0.056 publications per faculty-year, we would expect to observe negative values between -0.050 and -0.040, yet we observed $+0.022$ ($p < 0.02$ for all scenarios). The convergence of these approaches—each addressing different aspects of the possibility that contamination may be present—strengthens our conclusion that the null finding reflects a genuine absence of productivity impacts rather than measurement limitations or contamination-induced attenuation.