Sexual Misconduct and Scientific Production

Manuela R. Collis *
University of Toronto

This version: June 10, 2025 — Please click here for latest version

Abstract

Sexual misconduct at work remains a persistent challenge in contemporary organizations. Although a growing body of important work has documented the profound impact on survivors, fundamental questions about the broader organizational implications remain unexplored. This project investigates the potential effects of sexual misconduct in academia. Specifically, I ask: How do sexual misconduct incidents in university departments shape scientific production, collaboration patterns, and gender diversity? I draw on a novel database comprising over 1,000 documented cases of workplace sexual misconduct across multiple universities and disciplines and employ a coarse exact matching approach paired with a staggered difference-in-difference design. I collect information from 359 research-focused departments that experienced at least one sexual misconduct incident and around 5,000 control departments at research-focused universities.

THIS IS WORK IN PROGRESS
Please do not distribute or cite without the author's permission.
Thank you.

***Collis**: Rotman School of Management, 105 St. George Street, Toronto, ON M5S 3E6, Canada (email: manuela.collis@rotman.utoronto.ca).

This study was pre-registered on May 10, 2025, and can be accessed under https://osf.io/zx3gu once the embargo has expired.

Introduction

Sexual misconduct at the workplace is widespread and has complex and lasting consequences for affected individuals. 59% 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). Psychologists and organizational behavior scholars have documented that sexual misconduct at the workplace reduces job satisfaction, increases turnover, and worsens psychological well-being, such as anxiety, stress disorder, or depression (Schneider, Swan and Fitzgerald, 1997; Cortina and Berdahl, 2008; Fitzgerald and Cortina, 2018). Further research in economics provides evidence that sexual misconduct in the workplace drives women away from organizations (Folke and Rickne, 2022; Adams-Prassl et al., 2024; Collis and Van Effenterre, 2024; Batut, Coly and Schneider-Strawczynski, n.d.).

Today, many companies offer sexual harassment training in response. In fact, in states such as New York, Massachusetts, or California, harassment training is mandatory for most companies. However, non-disclosure agreements are regularly used and perpetrators are often not fired (Adams-Prassl et al., 2024; Barmes, 2023). What explains this firm behavior?

This paper offers productivity effects as an explanation. I find that the *occurrence* of sexual misconduct does not affect overall productivity. *Firing*, however, affects productivity negatively.

To study the relationship between sexual misconduct and productivity, I turn to university departments. This setting offers several distinct advantages. First, its main output, namely scientific publications, is of positive societal value and thus a meaningful outcome. Second, scientific publications are an established and rich measure of productivity. Third, research has become more collaborative and diverse, providing a suitable context to study this question.

In this paper, I ask *Do sexual misconduct incidents in university departments affect scientific production?*. To answer this question, I build a panel dataset for over 5,000 university departments across nine disciplines between 1980 and 2024. To construct the treated departments, I use a dataset of all sexual misconduct cases at universities in the United States that have been publicized between 1980 and 2024. I use the universe of universities and randomly select nine departments per university as the comparison university departments. To meaningfully study the relationship between sexual misconduct and productivity, I restrict my sample to *research-focused* universities where the perpetrator was a *faculty*, resulting in 359 treated and 5,076 control university departments.

I hand-collect over 3,000 lawsuits, news articles, or proprietary documents shared through online document libraries which allow me to establish a life cycle of a sexual misconduct case with three crucial points of time: the year when sexual misconduct starts to occur, the year when reporting starts to happen, and the year when the accused faculty leaves the department. I use these three points of time as three different "treatment" timings to establish the causal relationship of (1) the occurrence of sexual misconduct, (2) the public reporting of sexual misconduct, (3) the departure of the faculty facing sexual misconduct allegations. This rich information also allows me to capture contextual variation in each case that allows me to investigate how contextual factors shape productivity.

We have to assume that the publicized misconduct incidents are not random. To establish causality, I pair Coarsened Exact Matching (CEM) with a staggered difference-in-differences event study design. University departments are matched on five covariates: discrimination and harassment state laws, whether the university is public or private, discipline, research intensity of the university department, department size.

I provide causal evidence that *public reporting* of sexual misconduct incidents hurts university departments' overall productivity, but that the *occurrence* of sexual misconduct does not. The average number of publications per department member remains indistinguishable for departments where sexual misconduct occurred compared to control departments. The public reporting of sexual misconduct reduces the average number of publications per department member by 0.1 (p-value < 0.05). For a department of 20 faculty members, that would be two publications per year, for example. It has almost twice the impact on men compared to women. However, there seems to be more variation among men. Two possible explanations for this gender difference could be that either women leave the department or that the publicizing of sexual misconduct affects certain types of men differently. This is subject to further analysis. Next steps also include the examination of faculty who decide to co-author with the accused faculty and those who do not, looking at more outcome measures such as citations, journal impact factor, and direction of science. Lastly, I will take into account contextual factors, such as the role of the status of the targeted individual, the speed of institutional processes, among others.

This paper contributes to two streams of research. To my knowledge, it is the first study that provides causal evidence on the impact of sexual misconduct on broader organizational outcomes such as productivity. Second, this paper contributes to the literature on inequality in science and innovation.

Taken together, this paper is motivated by the puzzle that sexual misconduct hurts targeted individuals and contributes to gender inequality in the labor market, and yet organizations seem to spend substantial resources on cover-ups. The explanation for this paradox appears to be that the *occurrence* of sexual misconduct in a university department does not noticeably impact productivity, while the *public reporting* of it does.

1 Empirical Design, Data, and Estimation

To study the impact of sexual misconduct incidents on organizational outcomes in academia, I construct a dataset that looks at the productivity outcomes of university departments.[1] The departments are listed in Table A. Treated departments have experienced a sexual misconduct case that became public.

[1]I define departments by following Clarivate's Core Collection containing 256 categories (https://mjl.clarivate.com/help-center and supplement these with SCImago categories when necessary (https://www.scimagojr.com/journalrank.php). Note, that for the purpose 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.

1.1 Treated Research Departments

To construct the dataset of treated research departments, I leverage a dataset of all sexual misconduct incidents in academia that have been publicized. I use data from the Academic Sexual Misconduct Database (ASMD) Libarkin (2024) to define a list of academics who have been linked to an exposed sexual misconduct incident. The database accessed on November 19, 2024 contains 1,294 incidents[2]. These are incidents which became public either through a lawsuit or media reporting and include resolved and ongoing incidents. Note, that when I say *publicized*, then I mean the newspapers have reported about the start of the investigation or lawsuit, any updates, or the conclusion of such. The incidents involve faculty members, administrators, coaches, and other staff. These cases have concluded between 1980 and 2024 (see Figure 1). For the purpose of this study, I focus on *resolved incidents* at *research-focused* institutions committed by *faculty*, which reduces the sample to 359 incidents.[3]

To construct the outcome variables, I use the institution and department information, at which the respective incidents occurred and collect the publications, citation information, reference information, and abstracts for each institution x department pair. I collect publication data from OpenAlex. OpenAlex is an open-source database providing access to over 200 million publication records by combining databases such as Microsoft Academic Graph (MAG), Crossref, ORCID, Unpaywall, Pubmed, Pubmed Central, ISSN, and others (Priem, Piwowar and Orr, 2022).[4] OpenAlex prides itself on a superior linkage of information compared to alternative data providers such as Web of Science, Lens.org, or Dimensions.

I manually assign the department by examining each of the 359 incidents in the dataset. I assign the department based on four pieces of information: Libarkin's assignment of discipline,

[2]Note that 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 Carnegie Classifications as one of these three categories: Research 1 - Very High Research Spending and Doctorate Production, Research 2 - High Research Spending and Doctorate Production, Research Colleges and University. I conducted a manual search and classification for the 57 universities without a Carnegie Classification. This removes 575 universities which are mostly technical colleges or U.S. Military/Navy Academies. Among the research-active universities, I remove ongoing cases (173) and cases for which both the individual and the discipline were unknown (83). Lastly, I remove all incidents which were committed by a coach, administrator, or unknown (126).

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

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 list of departments consists of topic categories constructed by Clarivate and SJR where each topic category is linked to a corresponding journal list.[6]. There are a total of 527 distinct topic categories which are listed in Section A. My dataset consists of 142 unique topic categories.

The 359 incidents took place at 139 unique institutions, one institution is anonymous. All 139 institutions could be matched to OpenAlex.[7] Of the 359 incidents, I can link 347 accused and publicly exposed faculty to an OpenAlex profile. This contains four names and eight incidents where we have enough information to assign a department and university but not enough information to identify the accused individual.[8]

1.2 Identification Strategy and the Construction of Control Research Departments

Given my interest in the effect of a sexual misconduct incident on scientific production, my empirical strategy is focused on changes in published research output after the incident (A) occurred, (B) was reported on by the media, and (C) led to the departure of the faculty facing sexual misconduct allegations, respectively, relative to before the incident occured, was reported on by the media, or led to the departure of the faculty facing sexual misconduct

[5] I hired research assistants who have experience in the life sciences and assigned the individuals in the social sciences myself.

[6] The SJR topic category "Economics and Econometrics"'s corresponding journal list, for example, consists of 166 journals ranging from "American Economic Review" to "Economics of Energy and Environmental Policy" or "NBER Macroeconomics Annual book series"

[7] I employed an automatic matching by institution name. Each institution that did not return a 1:1 match was matched manually where I search for the correct match if no potential match has been returned by the search algorithm and remove false positives. Note that I also remove universities which are abroad. For example, Duke University (https://openalex.org/I170897317) has a partner university with Wuhan University. That partner university is called Duke Kunshan University and is located in China (https://openalex.org/I4210159968). I remove that university since it is not located in the US and most likely unaware of inter-personal dynamics at the department of interest.

[8] This means all 12 cases can be used to estimate the total effect but they cannot be used to estimate peer effects.

allegations, respectively.[9] 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 coarse exact matching approach described below.

To construct the control research departments, I use the universe of U.S. based research-focused universities over the 1980-2022 time period. To do that, I first construct a panel dataset of the IPEDS data between 1980 and 2023 and then construct a list of unique universities. 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 according to the Higher Education Act of 1965. Note, that 2023 is the most recent data available.

I keep universities which have been categorized as research-active [10] according to the Carnegie Classification at least once in the history of Carnegie Classification. For that, I construct a panel dataset of the Carnegie Classifications between 1974 and 2023. I keep all university names which have ever been designated as a "Research University" during this time period. I then link the 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. I then link these universities to OpenAlex. There are a total of 1152 research universities of which 582 could not be matched with OpenAlex. I inspected all 582 universities. I was unable to manually match them to OpenAlex. A closer look of a subsample shows that they are not research-intensive universities. This leaves us with 564 control universities.

To construct control research *departments*, I assign nine randomly selected departments, where I select a department for each of the nine disciplines defined by Clarivate's Core Collection. This will construct 5,076 total control departments, 564 control departments per discipline. Note, since this is the universe of research-active universities x discipline pairs, a

[9]Note that in approximately 65% of the reported cases, the accused faculty left the institution.

[10]Defined by Carnegie Classifications as one of these three categories: Research 1 - Very High Research Spending and Doctorate Production, Research 2 - High Research Spending and Doctorate Production, Research Colleges and University.

subset of these university x department pairs is, by definition, treated. Instead of removing them from the control set, I will account for that using a staggered treatment design. This will be discussed in Section 2.

1.3 Variable Construction

Outcome Variables

Publication Count I measure department x year productivity by counting the number of publications published every year by a given university department. The publication data will be assembled at the department x year x publication level and then aggregated at the department x year or department x year x subgroup (for example, women faculty) level. To capture the productivity of all department members equally, I will count within-department co-authored publications multiple times. For example, if two department members co-author together, their publication will count as two publications.

Citations For each publication, I count citations received annually from 2014 through 2024. In contrast to the publication count, I will count each publication's citations only once, regardless of how many co-authors are affiliated with the department. This measure captures how frequently scholars acknowledge or build upon their peers' scholarly contributions, which is a key indicator of research impact rather than simply measuring department-level productivity directly. To extend this coverage to years prior to 2014, I plan to use the Clarivate Web of Science data.

Journal Impact Factor To measure the quality of the work produced, I will use the database by Marx and Fuegi (2021), which covers journals from 1811 to 2019. To extend this coverage until 2024, I am currently working on constructing an impact factor measure that covers 1975 to 2024. For that, I plan to use the Clarivate Journal Impact Factor (JIF), which is derived from the Clarivate Citation Report and calculates the number of citations per publication a journal received in a given year for publications which were published the prior year. The JFI has been published every year since 1975 (Clarivate and Mangan, 2022) and made data from 1997 until 2023, including, available to me.

Direction of Science I measure the direction of science in the form of semantic similarity of the corpus of abstracts across years for every given department. I compute the semantic similarity between abstract corpora across consecutive years for each academic department. This temporal analysis employs large language model (LLM) embeddings, which represent state-of-the-art semantic representation capabilities in natural language processing. By measuring cosine similarities between year-specific embedding centroids, this approach captures subtle shifts in research focus, terminology, and conceptual frameworks that characterize the trajectory of scientific inquiry within disciplinary boundaries.

Gender I will infer the department members' *gender* based on their first and, if available, middle name using genderize.io. I will use a statistical probability cut-off point of 90% for either the first or middle names or 75% for first and middle names jointly.

Women representation I will count the number of publishing women faculty for any given university department and year.

Treatment Indicators I collect all publicly available materials on the set of publicly known sexual misconduct cases in the U.S. Academy (Libarkin, 2024). This includes news articles, lawsuit documents, proprietary documents posted on digital document libraries, and information shared in the Academic Sexual Misconduct Database. Based on this rich set of information, I code up the treatment indicators for *occurrence of a sexual misconduct*, *media reporting of a sexual misconduct incident*, and *departure of a faculty member facing sexual misconduct allegations* at the year level.

Heterogeneity in Treatment Effect

Spillover Effects by Gender I will investigate whether the impact of sexual misconduct is different for men and for women department members, where I will infer their *gender* based on their first name using genderize.io.

Spillover by Co-Authorship Status To understand the *spillover effects* on the department and the colleagues, I will calculate the outcome measure for the treatment group in three ways: (1) *Total Spillover Effect*, which is the outcome measure at the department level, (2) *Co-authors*, which looks at the scientific production of department members who co-authored with the accused faculty, and (3) *Non-Co-authors*, which looks at the scientific production of department members who don't collaborate with the accused faculty. This decomposition of the total effect allows me to disentangle the spillover effect the accused faculty had on department members who co-authored with them versus not.

Counterfactual for Accused Faculty To measure how the effect varies for co-authors and non-coauthors, I need to construct a counterfactual for the accused faculty in the control departments. I will do that by ranking faculty according to their productivity for each department x year. I will take the rank of the accused faculty from the treated department and select the faculty with the same rank in the control department as our counterfactual control faculty.

Covariates To assign a research discipline to each department, I use the nine primary disciplines 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.

To determine whether the university is a public or a private university, I use IPEDS' institutional datasets. I am able to 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 the Shanghai Ranking'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). The ranking is ideal for the purpose of this research because it ranks universities based on research output and quality, faculty quality, and research collaborations.[11]. In contrast, U.S. News Education Ranking focuses on research output *and* student experience, and the Carnegie Classifications are based on the number of doctoral degrees granted and the amount of research expenditure. I assign the GRAS at the university x 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 state-level legal framework deviate from federal-law related to four categories: discrimination, workplace hostility, sexual harassment, and retaliation. To do so, I hand-collect the extent of legal deviations for each category.

Department size is determined by the number of publishing researchers in a given year.

2 Empirical Strategy

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

Departments experience the three "treatments" — misconduct, public exposure, and exiting of the accused faculty — at different points in time. The treatment effect is identified by

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

comparing outcomes in treated departments before and after exposure, relative to matched control departments that have not yet been treated or will never be 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 are never treated.

The treatment in my setting is not randomly assigned. For a case to become public, it requires either involved parties that are willing to file a lawsuit or go public or local or national news media that is motivated and able to obtain evidence. This is more likely the case for sexual misconduct cases that are more severe or persistent and take place at public universities where freedom of information requests can be filed, for example. To account for potential differences in observed covariates, I use a multi-stage coarse exact matching approach.

This multi-stage coarse exact matching 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 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 fixed effects (ψ_u) in all specifications provides additional protection against potential confounding from institution-level factors. By differencing out all time-invariant university 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 fixed effects ensure that our estimates reflect the causal impact of misconduct at the department level rather than capturing university-level heterogeneity.

My baseline model will be a simple difference-in-differnce estimator which I will bild out into a staggered Difference-in-Difference design (oftentimes called two-way fixed effects (TWFE)). Recent methodological developments allow me to address empirical challenges associated with staggered Difference-in-Difference 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 TWFE 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 where some effects receive negative weights, potentially leading to estimates with incorrect signs or magnitudes. I probe robustness to my results using the Sun and Abraham (2021) estimator which handles staggered Difference-in-Difference designs with treatment effect heterogeneity across cohorts and allows for matching weights.[12]

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

2.1 Stage 1: Coarsened Exact Matching (CEM)

I employ a multi-stage approach that combines preprocessing through Coarsened Exact Matching (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 based on pre-specified 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). Note that for simplicity, I use *publication* as the outcome variable Y moving forward. But the analyses described below will also be conducted for the other outcomes of interest (citations, journal impact factor, and direction of science).

[12] I will also report the estimators from Callaway and Sant'Anna (2021) and de Chaisemartin and D'Haultfœuille (2020) but since they don't allow for weights, they are not my preferred specification.

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. The matching covariates are selected to address the fundamental challenge of selection bias in observational studies of sexual misconduct. To satisfy the conditional independence assumption necessary for causal inference, I identify and match on covariates that simultaneously predict both the likelihood of experiencing sexual misconduct incidents (treatment assignment) and departmental research productivity (outcome). This approach helps isolate the causal effect by creating balanced comparison groups that differ primarily in their treatment status rather than in other confounding characteristics.

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

Legal Framework at the State Level. Institution's behavior is likely influenced by the law they are governed by. To account for that, I construct a categorical variable at the state level that captures deviation of state-level legal framework compared to federal-law related to discrimination, workplace hostility, sexual harassment, and retaliation. To do so, I hand-collect the legal deviations.

Research Intensity. The research intensity at a given university is assigned using the 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).

This matching approach ensures that treatment and control departments share similar characteristics across dimensions that jointly determine both the probability of experiencing observed sexual misconduct incidents and departmental research productivity. By creating well-balanced comparison groups through CEM before applying regression analysis, I reduce model dependence and strengthen causal identification (Ho et al., 2007).

Stage 2: Staggered Difference-in-Difference 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:

$$Y_{iut} = \alpha_i + \lambda_t + \psi_u + \beta_1 Treat_i + \beta_2 Post_{it} + \tau(Treat_i \times Post_{it}) + \gamma DeptSize_{it} + \varepsilon_{iut} \tag{1}$$

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

Standard errors are clustered at the university x department level, and control university x 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. The inclusion of university fixed effects ψ_u allows me to difference out any institution-level effects while identifying the department-level impact of sexual misconduct incidents. This approach controls for all time-invariant university characteristics that might confound the relationship between misconduct and departmental outcomes.

I will report the average treatment effect over the five post-treatment periods for which I will take the simple average of τ (Equation 3) with a standard error of the mean of the post-treatment coefficients. I will then continue with reporting the full event study in standard event-study plots.

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

Analysis 2: Heterogeneity Analysis by Gender and Co-Authorship Status Continuing with the matched sample, I investigate how the impact of misconduct incidents varies across different demographic groups within departments. I analyze:

1. **Gender**: Separate analyses for publications by women and men department members.

2. **Co-Authorship Status**: Separate analyses for department members who are co-authors of the accused faculty versus department members who are *not* co-authors of the accused faculty.

For each subgroup analysis, I estimate:

$$Y_{iut}^{g} = \alpha_i + \lambda_t + \psi_u + \beta_1 Treat_i + \beta_2 Post_{it} + \tau^{g}(Treat_i \times Post_{it}) + \gamma DeptSize_{it} + \varepsilon_{idt} \tag{4}$$

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

2.2 Staggered Event Study Design

To investigate the time-dynamic of the treatment effect, I estimate an event-study specification:

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

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 since it allows for matching weights and takes into account heterogeneous treatment effects over time.

3 Results

To start, I want to contextualize the sexual misconduct cases that take place in research-focused universities and are committed by faculty. These misconduct cases take place in all disciplines of academia, from Arts & Humanities (34.8% of all cases) and Social & Behavioral Sciences (26.2%) to Engineering, Computer & Technology (4.2%) or Business (1.1%). The cases also range 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. Note, that these categories are not mutually exclusive and that several cases involve multiple types of sexual misconduct. These harmful actions are mostly taken by tenured faculty (in 67% of the cases)[13] In over half of the cases, we are dealing with a repeated offender. Specifically, 12.5% of the cases, the accused faculty receives two allegations. In 39.8% of the cases, the accused faculty commits sexual misconduct three or more times.

The Coarsened Exact Matching (CEM) reduces the differences between the treated and the matched control university departments noticably from a multivariate L1 distance of 0.71 to 0.5 as shown in Table 2. When we look at the five matching variables, we can see that there is some imbalance for disciplines and matching on the department size. Further iterations of this paper will explore options to improve on that. When comparing the mean of the main outcome variables in Table 1, it is perhaps worth pointing out that the average number of publications per department member at baseline is larger for the control departments than for the treated departments both before (2.18 vs 1.53) and after (2.98 vs 1.51) matching.

[13] Note, that this number is likely slightly upward biased since in some cases, it was unclear whether the stated rank was at time of misconduct or time of reporting. A next round of coding will make this more precise.

The staggered difference-in-difference event study estimates in Table 3 are based on the CEM weights and use the Sun and Abraham (2021) estimators. I find that the *occurrence* of sexual misconduct in a university department does not impact a department's productivity. And this remains true when we look at women and when we look at men. These results contrast the productivity outcomes after sexual misconduct has been publicly reported on or after an accused faculty got fired. In those cases, productivity starts to decline significantly two years from the event, but seems to rebounce. Overall, this means for three years, a university department looses out about one publication for every ten department members. This effect is larger for men than for women. Men loose out on up to 0.17 publication for a given man department member. However, the effect is more precisely estimated for women. Since the sample size for men is higher than for women, this may suggest that some men are impacted more than others. To explore this further, I will conduct additional anaysis.

4 Figures and Tables

Figures

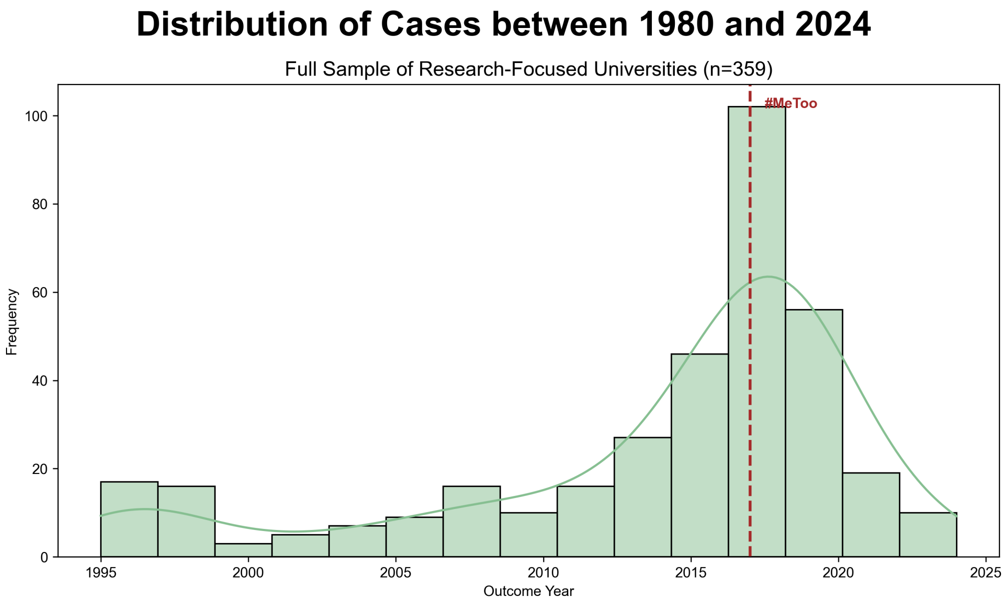

Figure 1: Distribution of Cases used for this Study

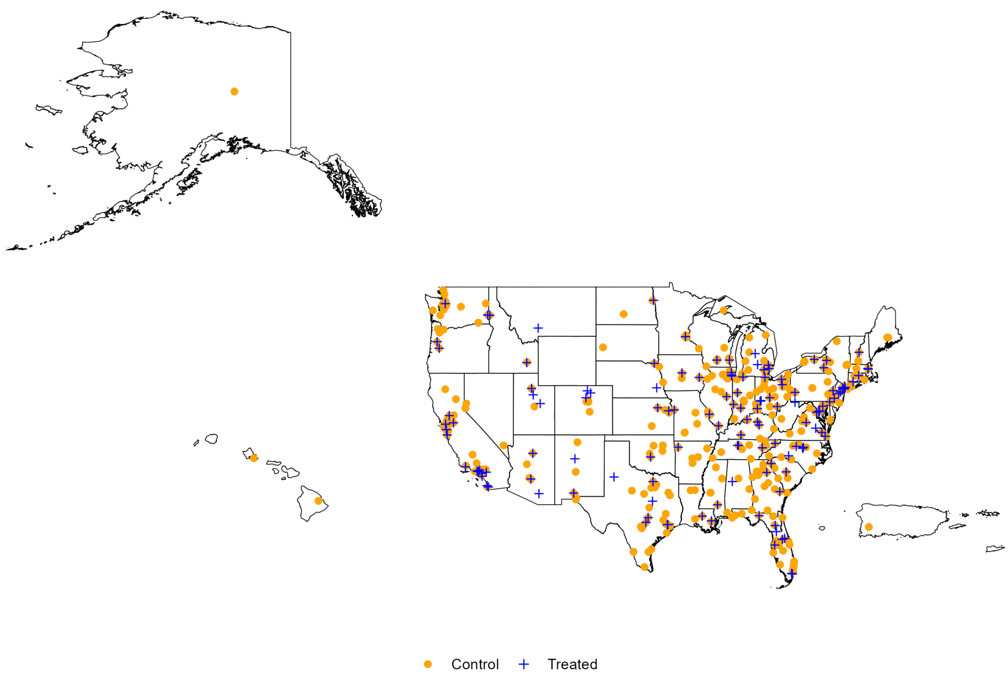

Figure 2: Treated and Control Universities - Treated and Control

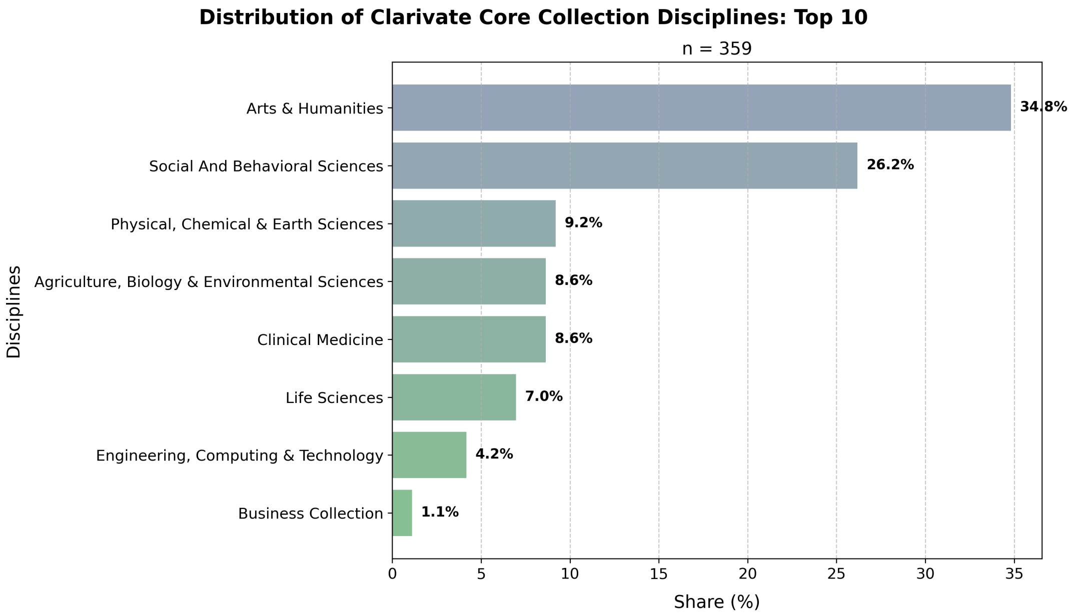

Figure 3: Distribution of Sexual Misconduct Cases Across Disciplines

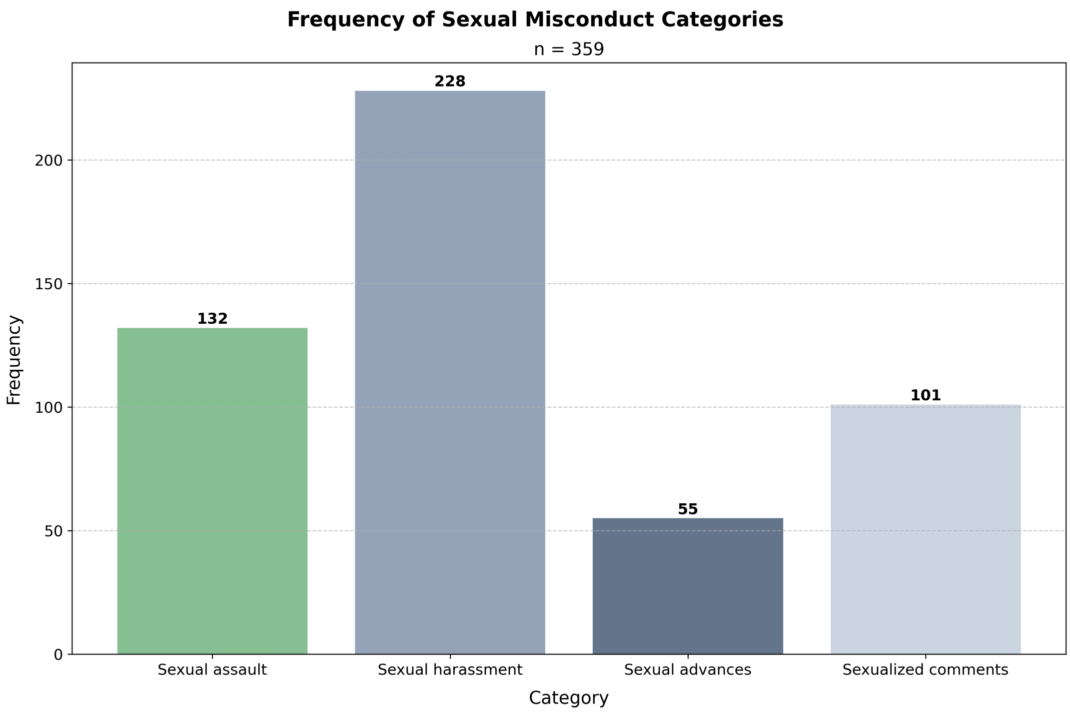

Figure 4: The Type of Misconduct Varies

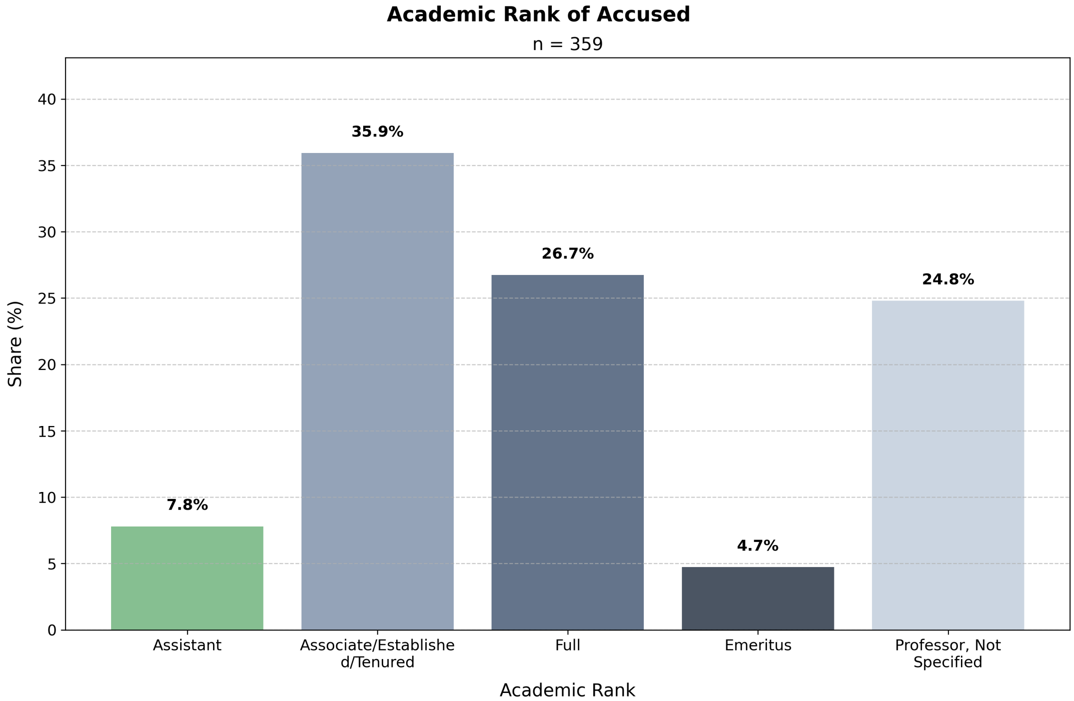

Figure 5: Majority of Accused Faculty Are Tenured

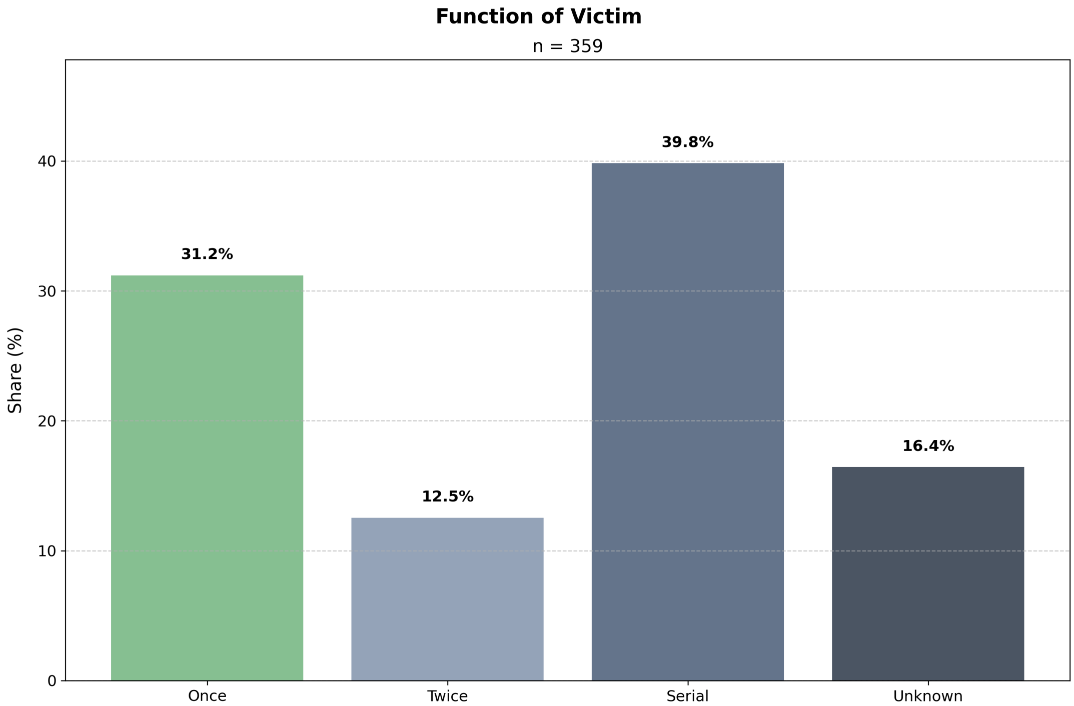

Figure 6: Majority of Accused Faculty Are Accused More Than Once

Tables

	Before Matching		After Matching		
Variable	**Control**	**Treated**	**Control**	**Treated**	**SMD**
Publications per Capita Total	2.18	1.53	2.98	1.51	0.549
	(2.65)	(1.11)	(3.53)	(1.02)	
Sample Size (N)	50,188	210	4,759	202	
Publications per Capita Women	2.28	1.51	2.99	1.48	0.579
	(2.73)	(1.11)	(3.47)	(0.94)	
Sample Size (N)	33,632	188	3,492	181	
Publications per Capita Men	2.33	1.64	3.14	1.59	0.528
	(2.94)	(1.42)	(3.76)	(1.28)	
Sample Size (N)	44,991	198	4,268	190	

Note: Values shown as Mean (SD). Control group statistics after matching use analytic weights. SMD = Standardized Mean Difference.

Table 1: Balance Statistics Before and After Matching

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

	Occurrence of Sexual Misconduct			Public Reporting of Sexual misconduct			Departure of Accused Faculty		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Total	Women	Men	Total	Women	Men	Total	Women	Men
5 Year Average TE	-0.006	0.018†	-0.032	-0.095*	-0.06**	-0.11†	-0.093*	-0.058*	-0.117†
	(0.021)	(0.024)	(0.023)	(0.048)	(0.016)	(0.025)	(0.045)	(0.026)	(0.061)
6 periods before treatment	-0.011	-0.-023	-0.005	-0.035	-0.037	-0.041	0.002	-0.003	-0.018
	(0.033)	(0.037)	(0.047)	(0.032)	(0.032)	(0.044)	(0.038)	(0.04)	(0.047)
5 periods before treatment	0.014	0.030	0.032	-0.017	-0.048	-0.030	-0.017	-0.053	-0.027
	(0.035)	(0.51)	(0.049)	(0.026)	(0.034)	(0.035)	(0.029)	(0.043)	(0.043)
4 periods before treatment	-0.022	-0.034	-0.026	-0.011	-0.050	0.002	0.001	-0.031	-0.016
	(0.032)	(0.031)	(0.045)	(0.032)	(0.026)	(0.49)	(0.027)	(0.030)	(0.041)
3 periods before treatment	0.023	-0.007	0.020	-0.011	-0.017	-0.013	-0.009	-0.016	-0.027
	(0.032)	(0.033)	(0.043)	(0.027)	(0.034)	(0.032)	(0.023)	(0.037)	(0.032)
2 periods before treatment	0.035	0.032	0.023	-0.063	-0.012	-0.091**	-0.027	0.005	-0.063
	(0.031)	(0.029)	(0.040)	(0.029)	(0.036)	(0.033)	(0.031)	(0.039)	(0.036)
1 periods before treatment	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)
Event period (t=0)	-0.012	0.008	-0.033	-0.043	-0.018	-0.044	-0.063†	-0.053	-0.063 †
	(0.025)	(0.031)	(0.033)	(0.034)	(0.033)	(0.04)	(0.033)	(0.036)	(0.037)
1 periods after treatment	-0.021	0.009	-0.011	-0.091†	-0.048	-0.087	-0.099*	-0.044	-0.108*
	(0.025)	(0.036)	(0.033)	(0.047)	(0.037)	(0.054)	(0.042)	(0.033)	(0.048)
2 periods after treatment	0.055	0.093	-0.008	-0.113*	-0.096***	-0.125*	-0.0121**	-0.041	-0.136*
	(0.070)	(0.076)	(0.037)	(0.049)	(0.031)	(0.062)	(0.046)	(0.042)	(0.057)
3 periods after treatment	-0.037	-0.012	-0.049	-0.132*	-0.099***	-0.157†	-0.133*	-0.086	-0.171*
	(0.026)	(0.030)	(0.031)	(0.061)	(0.029)	(0.085)	(0.058)	(0.034)	(0.086)
4 periods after treatment	-0.016	-0.006	-0.059	-0.094	-0.059***	-0.159	-0.049	-0.071*	-0.108
	(0.034)	(0.029)	(0.039)	(0.075)	(0.046)	(0.104)	(0.080)	(0.033)	(0.109)
Observations	53,841	29,623	35,880	39,484	29,623	38,240	39,484	36,616	35,880
R-squared	0.9385	0.9448	0.8893	0.9361	0.9448	0.8311	0.9361	0.6479	0.8889

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Standard errors clustered at university x department level. Fixed effects for university x department and years. Column 3 and 7 (Poisson models) report Pseudo R-squared in place of R-squared. *NR:* Sun & Abraham does not report on covariates. Column 8 uses instrumental variables. Column 9 uses propensity score matching.

Table 3: Effect of Sexual Misconduct on Publication per Department Member

References

Adams-Prassl, Abi, Kristiina Huttunen, Emily Nix, and Ning Zhang. 2024. "Violence Against Women at Work." *Q. J. Econ.*

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

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

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

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

Clarivate, and Rachel Mangan. 2022. "Journal Citation Reports Essentials 28_11_2022."

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.

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

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.

Folke, Olle, and Johanna Rickne. 2022. "Sexual Harassment and Gender Inequality in the Labor Market*." *Q. J. Econ.*, 137(4): 2163–2212.

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

Graf, Nikki. 2018. "Sexual Harassment at Work in the Era of #MeToo." *https://www.pewresearch.org/social-trends/2018/04/04/sexual-harassment-at-work-in-the-era-of-metoo/*, Accessed: 2025-4-28.

Ho, Daniel E, Kosuke Imai, Gary King, and Elizabeth A Stuart. 2007. "Matching as non-parametric preprocessing for reducing model dependence in parametric causal inference." *Polit. Anal.*, 15(3): 199–236.

Iacus, Stefano M, Gary King, and Giuseppe Porro. 2011. "Multivariate matching methods that are monotonic imbalance bounding." *J. Am. Stat. Assoc.*, 106(493): 345–361.

Iacus, Stefano M, Gary King, and Giuseppe Porro. 2012. "Causal inference without balance checking: Coarsened Exact Matching." *Polit. Anal.*, 20(1): 1–24.

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

Marx, Matt, and Aaron Fuegi. 2021. "Reliance on Science in Patenting."

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

Schneider, K T, S Swan, and L F Fitzgerald. 1997. "Job-related and psychological effects of sexual harassment in the workplace: empirical evidence from two organizations." *J. Appl. Psychol.*, 82(3): 401–415.

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

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

(For Online Publication)

Appendix to

Misconduct in Academia

Manuela R. Collis

List of Appendices

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