Allegations of Sexual Misconduct, Accused Scientists, and Their Research

Rainer Widmann (MPI-IC)
Michael E. Rose (MPI-IC)
Marina Chugunova (MPI-IC)

Discussion Paper No. 419

August 21, 2023

Collaborative Research Center Transregio 190 | www.rationality-and-competition.de
Ludwig-Maximilians-Universität München | Humboldt-Universität zu Berlin
Spokesperson: Prof. Georg Weizsäcker, Ph.D., Humboldt University Berlin, 10117 Berlin, Germany
info@rationality-and-competition.de

Allegations of Sexual Misconduct, Accused Scientists, and Their Research*

Rainer Widmann[1] ⓡ Michael E. Rose[1] ⓡ Marina Chugunova[1]

[1]Max Planck Institute for Innovation and Competition, Munich, Germany

July 2023

Abstract

Does the scientific community sanction sexual misconduct? Using a sample of scientists accused of sexual misconduct at US universities, we find that their prior work is cited less after allegations surface. The effect weakens with distance in the coauthorship network, indicating that researchers learn about allegations through their peers. Among the closest peers, male authors react more strongly, suggesting that they feel a greater need to disassociate themselves from the accused. In male-dominated fields, the effects on citations are more muted. Accused scientists are more likely to leave academic research, to move to non-university institutions, and to publish less.

Keywords: sexual misconduct, scientific community, scientific impact.
JEL-classification: J16, M14, I23, K4.

*rainer.widmann@ip.mpg.de; michael.rose@ip.mpg.de; marina.chugunova@ip.mpg.de (corresponding author). The author order is randomized. For able research assistance we thank Carolin Formella, Nishan Lin, Iliana Radeva, Cerensu Dertli, Oytun Tükenmez, Sophia Wetzler and Dennis Klein. For helpful comments and suggestions we thank Ayelet Gneezy, Uri Gneezy, Dietmar Harhoff, Matthew Higgins, Karin Hoisl, Benjamin Jones, Nicola Lacetera, Petra Moser, Wolfgang Luhan, and Henry Sauermann as well as the audiences of seminars and conferences. Chugunova and Widmann gratefully acknowledge support from the Deutsche Forschungsgemeinschaft through CRC TRR 190 (project number 280092119).

1 Introduction

Sexual misconduct remains a pervasive and pressing issue that requires immediate attention (Cantor et al. 2019; Feldblum and Lipnic 2016; Maestas et al. 2017). However, individuals confronted with allegations of misconduct face ethical decisions in which they must carefully weigh competing values such as upholding social norms versus advancing the goals of their organization or profession. These complex dilemma force individuals to grapple with difficult choices, including whether to separate the work of the accused from their personal actions, acknowledge the contributions of violators, or collaborate with individuals who have otherwise made significant contributions. In this article, we focus on academic research.

In science, the goal is to produce knowledge. To facilitate the prolificacy of the process, science is organized around a set of principles known as the “Mertonian norms” (Merton 1973). One tenet, among others, is that ideas are evaluated on their own merit, regardless of who created them. Yet, at the same time, science is also a social system, and the community of scientists relies on additional norms to create an inclusive environment. In this paper, we try to answer the question of whether the scientific community not only sanctions “bad science”, but also “bad citizenship”. We study the effect of sexual misconduct allegations on the scientific impact of the accused scientist’s body of work, their publication output, and their employment, providing the first systematic empirical evidence on the consequences of sexual misconduct allegations for the accused.

We construct a dataset of 212 scientists at research-intensive universities in the United States across all disciplines against whom allegations of sexual misconduct have been made public between 1998 and 2019. We require that the cases have been disclosed in a newspaper article, in a university investigation, or in a court report, and that there has been an action taken in response to the allegations that substantiates them (the accused was found guilty or admitted guilt, left the position - fired, resigned, retired, died - or settled). In our sample, allegations that are ultimately found to be false are likely to represent a very small number of cases.

By virtue of its construction, our sample focuses on incidents involving well-published, senior researchers and on incidents that are well covered by newspaper media. Since the publicity of sexual harassment allegations is a prerequisite for others to respond and for the allegations to become constitutive of public discourse, we believe this sample is the most appropriate for answering the question of whether the scientific community sanctions sexual misconduct. We acknowledge that our study focuses on the small "visible peak" of the largely invisible "iceberg phenomenon" of sexual misconduct at universities (Clancy et al. 2020). However, as it is impossible to collect administrative data on sexual misconduct, we view our approach of relying on publicly available documents as virtually without alternative.

We connect accused scientists to their publication profiles in the bibliographic database Scopus and track citations to prior publications of accused scientists over time to detect changes in the citing behavior of other researchers after the allegation became public. We compare the citation trajectories of their articles to the trajectories of other articles in the same journal issue, which capture the counterfactual outcomes in the absence of allegations. We further examine how adjustments in the citing behavior of other researchers depend on their own distance in the coauthorship network from the accused, on their gender and on the gender ratio of the field in which they are active.

We document an overall decline in citations to the prior of work of the accused scientist, driven by researchers who are close to the accused in the network of coauthors (distance 1 or 2, i.e. coauthors or coauthors of coauthors), while researchers who are distant do not appear to respond at all. Gender differences in citing behavior are limited to authors closest to the accused in the network, where male authors reduce their citations by about twice as much as their female counterparts. We show that the effect is limited to articles that are published in fields with a high proportion of female authors.

In addition, we aggregate information at the scientist level and examine changes in publication output, collaborations and affiliations. We match accused scientists to a set of observationally similar control scientists based on, among other characteristics, field, academic age, publications and coauthors. We find that accused scientists publish less, leave academic research at higher rates and are more likely to move to non-university research institutions after allegations become public. We find a decrease in collaboration with others, which does not appear to be gender specific and is largely explained by the overall decrease in publication output as the average number of coauthors per published paper does not decrease.

Our study relates to several strands of literature. First, we draw on the ideas and methods of studies on the “retraction penalty", which have documented that, after a retraction, citation losses spread to the researcher’s entire body of prior work (Azoulay et al. 2017; Jin et al. 2019; Lu et al. 2013). Such a penalty may be conceived as consistent with the Mertonian norms, as a retraction indicates a violation of professional norms. We show that a penalty of a similar magnitude applies to the contributions of scientists who have violated social norms.

Second, we advance the research on sexual misconduct by providing first systematic empirical evidence on the consequences for the accused. It has been documented that sexual harassment in the workplace is detrimental for the victim’s job satisfaction (Fitzgerald et al. 1988) and has long-term health effects (Chan et al. 2008). In addition, studies indicate substantial economic costs for employers, public institutions, and society as a whole: e.g. due to lower productivity, distorted occupational choices, and career attainment (Antecol and Cobb-Clark 2006; Chan et al. 2008; Cici et al. 2021; Folke and Rickne 2022; Laband and Lentz 1998; McLaughlin et al. 2017; National Academies of Sciences, Engineering, and Medicine 2018). Consistent with

Luo and Zhang (2022), who show that associates of Harvey Weinstein increased their hiring of female writers after the allegations broke, we find that peers who are connected to the accused in science tend to change their behavior the most in response to allegations.[1]

Third, our study contributes to the literature on gender disparities and culture within the scientific community. It has been shown that women are less likely to be mentioned as authors on joint projects (Ross et al. 2022), receive less credit for collaborative work (Sarsons et al. 2021) and are less likely to be cited in top journals and by men (Koffi 2021). How the scientific community deals with sexual misconduct will affect women's prospects of entering and advancing in academia. In addition to harming women directly, fear of sexual misconduct allegations may also distort collaboration between the sexes (Gertsberg 2022).

Finally, we relate to recent studies of punishment for workplace misconduct that focus on financial fraud (Egan et al. 2022) or violent assault (Adams-Prassl et al. 2022). We consider peer punishment in addition to punishment from the employer.

2 Data

Sexual misconduct allegations Our primary source is the Academic Sexual Misconduct Database (ASM) by Libarkin (2019), which contains allegations of sexual misconduct involving faculty and other university employees at U.S. universities.[2] It was populated by weekly Google

[1]Focusing on under-reporting of sexual harassment, Dahl and Knepper (2021) establish that a threat of retaliation is an important barrier and Cheng and Hsiaw (2022) suggest that raising awareness may alleviate under-reporting.

[2]We search other sources: (A) the WestLaw database of court records, following the search strategy of Cantalupo and Kidder (2018) and extending it to cover the years 2013-2020; (B) the Title IX database of Chronicle of Higher Education webpage; (C) the database of 'Daily Californian' of all Title IX investigations in the University of California system. We apply the inclusion criteria of the ASM to the other data sources. The search strategy by ASM has proven comprehensive with only 10 additional cases in the final sample stemming from other sources: 9 from the WestLaw and 1 from 'Daily Californian'.

searches beginning in 2016,[3] using a combination of sexual misconduct keywords and academic workplace keywords (see Table S1). This database offers a critical advance to the de facto absence of data on scientists accused of sexual misconduct and has been used in studies both in economics (Gertsberg 2022) and in other fields (Cantalupo and Kidder 2018; Espinoza and Hsiehchen 2020). To be included in the database, a document (either a newspaper article, a court report, or a university investigation report) must reveal that at least one of the following events had occurred: (1) An institution found a faculty guilty; (2) An institution fired a faculty in response to allegations; (3) There is an admission on the part of the accused; (4) The accused faculty resigned, retired, or died before an institution could complete an investigation; (5) A faculty settled with their accuser; (6) A court made a finding of legal fact that sexual harassment had occurred, with or without legal punishment (criminal) or payment (civil). We refer to the year in which this event occurred as the "outcome year" and assume that the effects materialize thereafter. We suppose that the event lends credibility to the allegations and makes them public.[4] We extract all cases that name research faculty at research-intensive universities (R1 or R2 according to the Carnegie classification) with an outcome year after 1998 and before 2019, inclusive. Our study period starts after the year 1997, when the first federal guidance to unify universities' handling of sexual misconduct was provided in the "letter" of the U.S. Office of Civil Rights (Education Office for Civil Rights 1997; Johnson 2015; Lucey et al. 2020) (see details in the appendix A.1). We find 305 accused scientists, each linked to one incident.[5] We connect this sample to our bibliometric data and retain all cases where the accused can be linked to at least one research-type journal article that predates the outcome year. This step reduces our sample to 212 accused scientists.

[3]When the database was created, the described search procedure was used for all prior years.

[4]In 81.6% of cases, the outcome year predates all newspaper mentions of the case. In 13.2% of cases, the earliest newspaper mention appears in the year immediately preceding the outcome year. As a robustness check, we consider the earliest mention of the case in a newspaper article to be the treatment date if it predates the outcome year.

[5]We identified 310 incidents that involved 305 scientists. For scientists with more than one incident, we consider the earliest.

Background Information We manually collect additional background information on our cases from public sources to better understand the characteristics of our sample. We search LexisNexis, a database of legal and journalistic documents that covers university newspapers, local newspapers and federal newspapers. We find that 209 of cases (98.6%) are reported on in at least one newspaper article. 47.2% of cases are covered in at least one article by one of the 33 largest U.S. newspapers by circulation or covered by Associated Press.[6] The year of the first mention of the allegations largely coincides with the outcome year as defined above (56.9%) or is separated by no more than one or two years (34.9%). We additionally search for the year in which the incident allegedly occurred, which is available for 139 cases.

We categorize the nature of the alleged incidents largely according to the categories suggested by Fitzgerald et al. (1995). Incidents in our sample range from inappropriate comments (13.7%) and unwelcome sexual advances (48.6%) to stalking (2.8%), sexual relations (22.6%) and rape (1.9%).[7] In addition, we categorize the immediate disciplinary consequences the accused faces. By increasing gravity, the repercussions range from no consequences (0.9%), mild administrative consequences (3.3%) and monetary consequences (4.3%) to serious administrative measures (18.9%) - e.g., probation or suspension - and resignation (60.4%).[8] We also find cases in which the accused is convicted by court (6.1%) or commits suicide (2.4%).

Because some of the inclusion criteria of the ASM do not unambiguously signify the finding of guilt or an admission of guilt (e.g., settlement), we conduct a supplementary review for 40 randomly chosen cases in our sample. We read in detail all coverage available in LexisNexis until we find a definite statement of guilt or that the allegations were found false. We find no

[6]The list of the newspapers is the combined list of the top 25 newspapers in the U.S. for the year 2007 by Audit Bureau Circulation and for the year 2021 by Alliance for Audited Media. See the appendix A.3 for the exact list.

[7]10.3% do not fall into these categories or are unknown. Many incidentsinvolve actions that fall into more than one category, in which case we assign the most severe category.

[8]In the case of multiple consequences, we assign the most severe; the order of severity is determined by the authors. The last category captures both, voluntary and involuntary resignations, which oftentimes can not be clearly differentiated (Fortney and Morris 2021). 3.8% could not be classified.

case in which the allegations were found false, which leads us to believe that the proportion in the overall sample is very small. This is in line with the literature on sexual harassment that finds that false reporting is rare (Lisak et al. 2010; McCann et al. 2018).

Publications We match accused scientists to their author profiles on Scopus and retrieve their publications and the citations received by those publications, starting in 1971 and ending in 2021. We use the service gender-api.com to assign a gender to all coauthors and all citing authors. We record whether an article appears in a journal in a "male-dominated" field, which we define as all ASJC ("All Science Journal Classification") 2-digit fields with a share of female authors below 20% in the year 2000. We determine the percentile rank of the journal within its research field based on the SCImago Journal Ranking. Finally, we use affiliations recorded on publications to infer employer transitions. We categorize affiliations by type (university vs. non-university) and assign a rank or rank range in the Shanghai University Ranking. Appendix B.1 presents more details.

Coauthorship network For each citation, we determine the minimum distance d between the authors of the citing article and the accused in the global coauthorship network. To abstract from adjustments in the network due to allegations, we fix the coauthorship network in the year before the outcome year. We assign the distance based on links that were formed before the year in which the citing article was published or before the outcome year, whichever comes first. We consider two authors to be connected in a given year if they have ever published an article as coauthors in the past. We remove "inactive" researchers from the network who stopped publishing at least two years earlier. We define d such that $d = 1$ refers to former coauthors of the accused, $d = 2$ refers to coauthors of former coauthors and so on. Authors who cannot be connected through the network are assigned the distance $d = +\infty$.

Control articles and control scientists In the first part of the analysis, we compare the citations received by articles (co)authored by accused scientists ("treated" articles) to a group of (up to) 10 randomly selected other research articles from the same journal issue, provided that the authors are not also accused ("control" articles). If the journal has no issues, we sample from the entire volume. By matching on journal issue, we ensure that the articles are published at the same time, belong, broadly speaking, to the same subfield of the discipline and are of comparable quality. We choose this strategy, as opposed to matching articles based on citation pre-trends as in (Lu et al. 2013), because it is transparent and allows us to examine potential anticipatory effects (and other pre-trends). We track citations to control articles and randomly select one author to determine the coauthorship distance to the authors of the citing articles.

In the second part of the analysis, we compare accused scientists to other observationally similar scientists. They are matched based on characteristics four years before the outcome year or one year before the year when incident allegedly occurred, whichever comes first, using the package 'sosia' by Rose and Baruffaldi (2020). We match the treated and control scientists based on their research area, years since first publication, number of publications, coauthors and citations, country of affiliation, institution type and gender. We find at least one suitable match for 158 accused scientists (i.e. 54 without a match, which we discard for this part of the analysis) and randomly select one.

Descriptives Our sample includes 212 accused scientists and their 10,697 prior publications. 96.75% of the accused scientists are male. We present sample statistics in Table 1. The average career age at treatment, that is, the number of years since the first publication, is 24.6 years, while the average number of publications per year prior to treatment is 1.91. About one-third of all articles in our sample are published in journals that rank in the top 10% in their field. The average publication age, i.e. the number of years since publication, is 13.6 years. Figure S1 in the appendix shows the distribution of accused scientists across all scientific fields. Figure S2 in the appendix shows the distribution of outcome years in our sample, which spans the years 1998 to 2019.

Table 1: Sample statistics for accused scientists and their body of prior work

	Mean	Std. dev.	p10	p25	p50	p75	p90
Scientist level (N=212)							
Year of First Publication	1989.38	13.05	1972	1980	1991	1999	2005
Career Age at Treatment	24.59	12.38	10	15	23	31	43
Publications per year	1.91	2.09	0.20	0.41	1.10	2.71	4.78
Coauthors per publ.	2.85	4.39	0	0.67	2.00	3.48	5.62
Article level (N=10,697)							
Year of Publication	2001.19	10.95	1986	1994	2003	2010	2014
Publication Age at Treatment	13.62	10.30	3	6	11	19	29
Citations per year	3.54	11.89	0	0.24	1.07	3.17	7.67
Journal-rank percentile	21.42	16.06	4	8	17	34	50

Notes: Statistics are based on all years prior to the outcome year.

3 Results

3.1 Effect on citations to the accused scientist's body of prior work

We examine the effect of sexual misconduct allegations on the number of citations to the accused scientist's body of prior work using a "difference-in-differences" approach. Focusing on articles that were published prior to the outcome year, the effect is identified by the relative difference in the change of citation rates between "treated" articles (i.e., those of the accused scientist) and control articles after the outcome year. This approach allows us to tease out the change in the citation behavior of other researchers, while holding the cited content constant.

We exclude self-citations by all authors.

We consider a regression model of the form

$$y_{i,t} = a_i + \tau_{i,t} D_i I_{t,\tau^*(i)} + \gamma_{t,j(i)} + \varepsilon_{i,t} \tag{1}$$

where $y_{i,t}$ is the citation count of publication i in year t, transformed by the inverse hyperbolic sine, a_i is a publication fixed effect and $\gamma_{t,j(i)}$ is a fixed effect for the number of citations in year t for articles published in journal issue $j(i)$. D_i indicates whether the publication was (co)authored by an accused scientist and $I_{t,\tau^*(i)}$ is an indicator that captures timing relative to the outcome year $\tau^*(i)$. $\tau_{i,t}$ is a time- and article-specific treatment coefficient that captures the effect of allegations. When we examine the citation effect by gender of the citing first author or by distance in the coauthorship network, our dependent variable counts only the citations received by the specified group of articles.

We use the estimation strategy proposed by Borusyak et al. (2021), which is robust when the treatment effect depends on the time of treatment. This is important in the present analysis, because attitudes toward sexual misconduct have changed over the years. Our main estimate is the unweighted average of all time- and article-specific treatment coefficients from the specification $I_{t,\tau^*(i)} = \{t \geq \tau^*(i)\}$, denoted $\bar{\tau}$. In addition, we estimate the averages of the treatment coefficients for each year relative to the outcome year, given by $\bar{\tau}_T = \frac{1}{N_T} \sum_{i,t} \{\tau_{i,t} | t = \tau^*(i) + T\}$. We convert all estimates into percent semi-elasticities. The standard errors are clustered at the level of the accused scientist (in the treatment group) or at the article level (control group). Additional details are included in appendix C.1.

In Figure 1, we present estimates of the change in citation rates for treated articles relative to the control articles over time. We aggregate the evidence across all incidents in our sample and track citations through 2021. We find evidence of diverging citation trajectories. Citations

Figure 1: Effect of allegations of sexual misconduct on the accused scientist's body of prior work

Notes: Citation losses per year for articles published before the outcome year, compared to the control group. Blue dotted lines indicate 95% confidence intervals. Citations to the accused scientist's prior work decline on average by -5.25% per year (p<0.001) relative to the control group.

to accused scientists' articles decline by about -4.10% relative to the control group in the first five years after the outcome year (p-values for individual years range between p<0.001 and p=0.019). The effect does not seem to diminish over time.

In Figure 2 Panel A, we present our main estimate of -5.25% of the average yearly decline in citations over the entire outcome period (t=-5.50, p<0.001), shown in gray.[9] Figure 2 Panel A shows, in addition to our main estimate, citation losses by gender of the first author of the citing article. In Figure 2 Panel B shows citation losses by distance in the coauthorship network between the accused scientist and the authors of the citing article, and interacts this distance with the gender of the first author of the article.

Because there is a pronounced gender asymmetry in sexual misconduct, with men primarily being perpetrators and women primarily victims (Cantor et al. 2019; National Academies of

[9] All p-values for t-tests presented in this paper refer to two-sided tests.

Sciences, Engineering, and Medicine 2018), we expect asymmetry in the responses of female and male authors. We find that, overall, citations by articles with a female first author decrease by about -3.95% (shown in Figure 2 Panel A, t=-3.74, p<0.001), while, perhaps surprisingly, citations by articles with a male first author decrease by about -6.01% (t=-6.16, p<0.001). The asymmetry in response is pronounced when we focus on articles at distance $d = 1$, i.e. articles that include coauthors of the accused in the author team.[10] Citations by articles of coauthors of the accused that have a male first author show a decrease of -38.53% (Figure 2 Panel B, t=-7.60, p<0.001). On the other hand, citations by articles of coauthors of the accused that have a female first author only reduce by -19.80% (t=-3.51, p<0.001). When we move to articles with distance $d = 2$, i.e. articles that include coauthors of coauthors of the accused, the difference in the reaction of female and male first authors largely disappears (-11.00%, p<0.001, t=-5.98 for male first authors vs. -8.63%, p<0.001, t=-3.45 for female first authors).

We interpret the stronger response of male citing authors who are very close to the accused in the peer network, compared to female citing authors, as them actively trying to dispel suspicions that they are complicit or that they knowingly tolerate sexual misconduct among their collaborators. Assuming that women do not face this suspicion, but are ultimately at least as concerned about sexual misconduct as their male colleagues, we can attribute about half of the response in the citation behavior of male first authors of articles at distance $d = 1$ to a perceived need to disassociate themselves from the accused. This motive seems to disappear for male authors at distance $d = 2$ (or higher), perhaps because higher degree coauthorship associations are (typically) not obvious to other researchers. In the Conclusion and Discussion section, we discuss other motives that may drive the citation responses of both male and female authors.

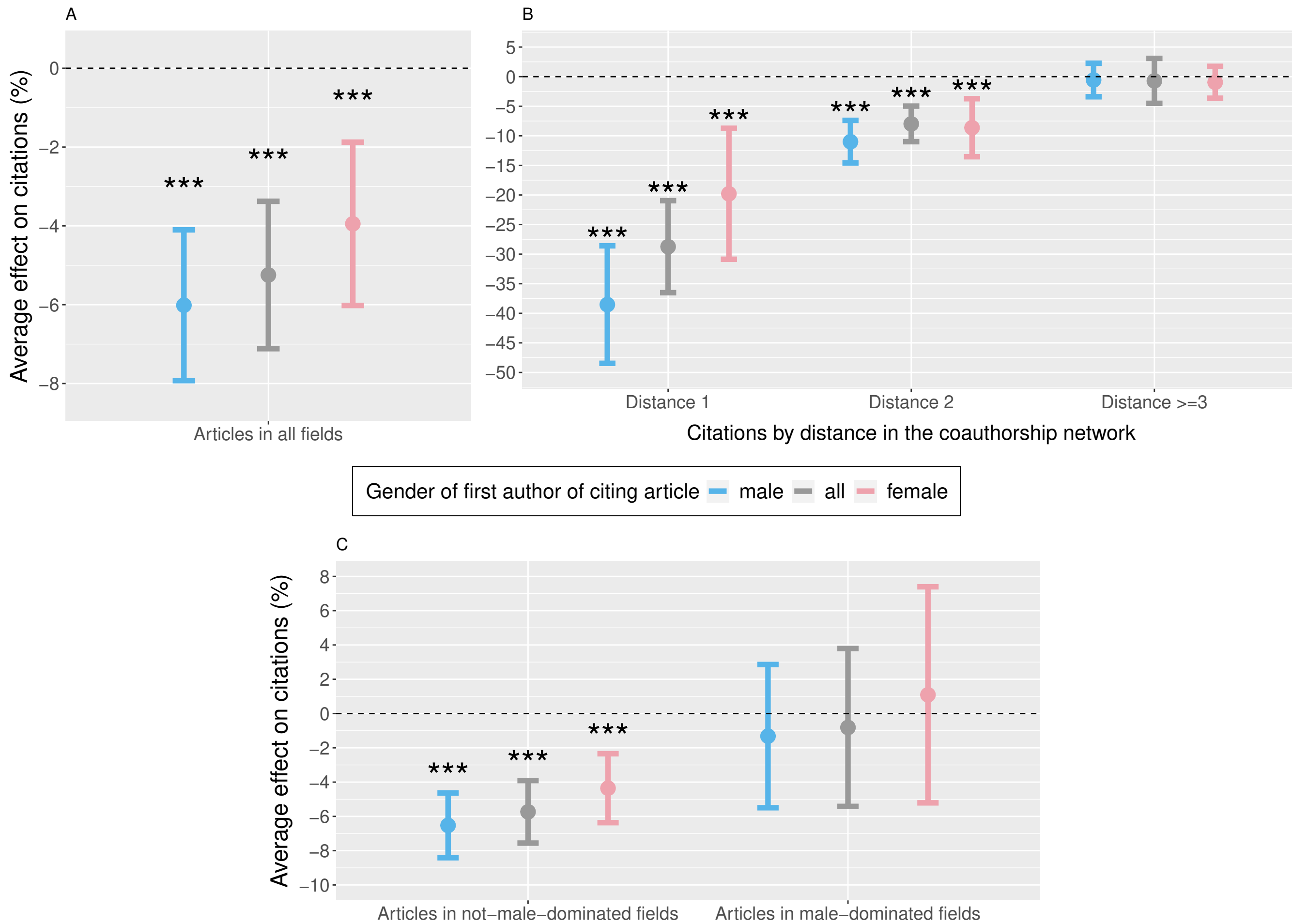
[10]To be clear, we consider the gender of the first author of the citing article, not the gender of the coauthors of the accused who are among the authors of the citing article. We do this to keep the classification of citing articles by gender constant across all panels in Figure 2. This approach implicitly assumes that the first author makes the decision whether or not to cite after observing the position in the coauthorship network of everyone in the author team relative to the accused.

How researchers learn about the allegations may depend on their position in the coauthorship network, too. Information may be transmitted either through peer networks or through other news sources. Assuming that the mode of learning is not gender specific, we posit that differences in the responses by distance in the network which are observed in both genders shed light on the relative importance of the two channels. Panel B shows that citation responses decrease sharply with distance in the coauthorship network, and that this pattern holds for both female and male citing authors. We find that citations by articles at distance $d = 1$ decline by -28.74% (shown in gray, t=-7.25, p<0.001), while citations by articles at distance $d = 2$ decline only by about -7.98% (t=-5.21, p<0.001). Citations from articles at distance $d = 3$ or more appear unaffected (-0.72%, t=-0.37, p=0.711). This pattern suggests that researchers learn about allegations through their peer network. The finding that distant authors do not respond further suggests that other news sources play a negligible role in spreading this information. We corroborate this point by examining how the effect depends on the publicity of the case in the newspaper media. When we split our sample[11] between incidents reported by one of the 33 largest newspapers in the U.S. (see appendix A.3) and those that are not, we find that they lead to very similar reductions in citation (-5.32%, t=-4.19, p<0.001 for the former group vs. -5.39%, t=-4.57, p<0.001 for the latter group).

A voluminous literature suggests that the gender composition of a field or a profession is associated with the prevalence of sexual misconduct (Dahl and Knepper 2021; Folke and Rickne 2022; Kabat-Farr and Cortina 2014; National Academies of Sciences, Engineering, and Medicine 2018). We examine differences in the strength of the citation response to allegations across different academic fields, which could be interpreted as a proxy for attitudes toward sexual misconduct. In Panel C, we examine whether the citation response is more muted for articles that are published in male-dominated fields, which we define as fields with a share of female authors

[11] Our estimation strategy, derived from Borusyak et al. (2021), requires that we examine heterogeneous effects for subsamples using split-sample estimates, rather interacting the effect with the respective subsample.

Figure 2: Citation effect by citing author group and by field of publication



Notes: Panel A shows average yearly citation losses after the outcome year when counting only citations by first authors of the specified gender. Panel B shows citation losses when counting only citations by authors of the specified distance in the coauthorship network, and by first authors of the specified gender. E.g. articles at distance $d = 1$ comprise coauthors of the accused. Panel C presents citation losses for articles that are published in fields with a female author-share of below 20% in the year 2000 (Engineering, Energy, Physics & Astronomy, Computer Science, Mathematics, Economics, Decision Science, Earth and Planetary Science, Materials Science, Chemical Engineering) and for articles in all other fields. In all panels, we track citations until 2021.

below 20% in the year 2000.[12] We acknowledge that the proportion of female researchers is correlated with other field characteristics (e.g. a field being in the natural sciences) that might also predict the strength of the citation response. When we split our sample accordingly, we find an estimated effect for articles published in male-dominated fields of -0.81% (shown in gray, t=-0.34, p=0.729). For the remaining fields, we find an effect of -5.73%(t=-6.16, p<0.001). Thus, we can reject effect sizes for male-dominated fields that are equal to or larger than those for not-male-dominated fields.[13] We examine the robustness of this finding to different definitions of male-dominated fields in appendix C.2.

When we differentiate by the gender of the citing author, we find that within both, male-dominated and not-male-dominated fields, male and female citing authors appear to behave similarly. This suggests that the field differences found are due to field-specific factors that affect male and female authors in the same way, such as the climate and culture of the field.

Last, we conduct an analysis that supports our claim that the decline in citations decline is due to the allegations themselves, and not due to a potential change in the accused scientist's research environment. For example, if the accused scientist moved to another institution after the allegations were made, the move could potentially inhibit activities that help disseminate his prior research. However, this alternative interpretation seems inconsistent with the fact that, when we restrict our sample to those cases where the accused scientist does not leave (either voluntarily or involuntarily), we still find a significant decrease in citations to his prior work (-5.06%, t=-4.03, p<0.001). Conversely, the effect appears only marginally larger for accused scientists who resign, are convicted, or who commit suicide (-5.34%, t=-4.68, p<0.001).

[12] According to this definition, the following fields are male-dominated: Engineering, Energy, Physics & Astronomy, Computer Science, Mathematics, Economics, Decision Science, Earth and Planetary Science, Materials Science, Chemical Engineering. We assign 2,217 treated articles (20.7%) that were (co)authored by 107 different accused scientists (out of 212) to male-dominated fields.

[13] In a one-sided test, the hypothesis $\bar{\tau}_{md} \leq$-5.73%, where $\bar{\tau}_{md}$ denotes the effect in male-dominated fields, is rejected with p=0.018.

Parallel trends and anticipatory effects It is conceivable that researchers are more likely to face allegations when their reputation is on the decline, but the opposite is also possible. We test whether there is evidence of differential trends in the period preceding the allegations. In the model described by equation (1), we test whether the averages of the pre-treatment coefficients for the 5 (10) years preceding the outcome year are jointly equal to zero, i.e. $H_0 : \bar{\tau}_{-5} = .. = \bar{\tau}_{-1} = 0$. We fail to reject the null-hypothesis at p=0.324 (p=0.281). We then test whether the averages of individual pre-treatment years equal zero and fail to reject the null-hypothesis in all cases, with p-values ranging between 0.186 and 0.863.

It is also possible that the allegations, or the publicly observable action that substantiate them, were anticipated by other researchers. We first examine the robustness of our results when we allow for anticipatory effects of varying time margins. If we allow the treatment year to precede the outcome year by one year, our estimate of the average yearly decline becomes -4.54% (t=-4.64, p<0.001), by two years it becomes -4.39% (t=-4.26, p<0.001), and by three years the estimate becomes -3.82% (t=-3.79, p<0.001). When we use the year of publication of the earliest newspaper article ever mentioning the allegations as the treatment date, provided it predates the outcome year (relevant for 39 cases), we find an estimated effect of -5.07% (t=-5.84, p<0.001). In another exercise, we use the year in which the incidents allegedly occurred as the treatment year, which is available for 139 incidents. We expect this date to precede any rumors about the incident. In this case, we find an estimated effect of -4.46% (t=-5.11, p<0.001).

3.2 Evidence on accused scientists' outcomes

We examine the effect of allegations on the number of publications and the number of collaborations of accused scientists in a regression model of the form

$$x_{s,t} = a_s + \tau_{s,t} D_s I_{t,\tau^*(s)} + \gamma_t + \eta_{f(s,t)} + \varepsilon_{s,t} \tag{2}$$

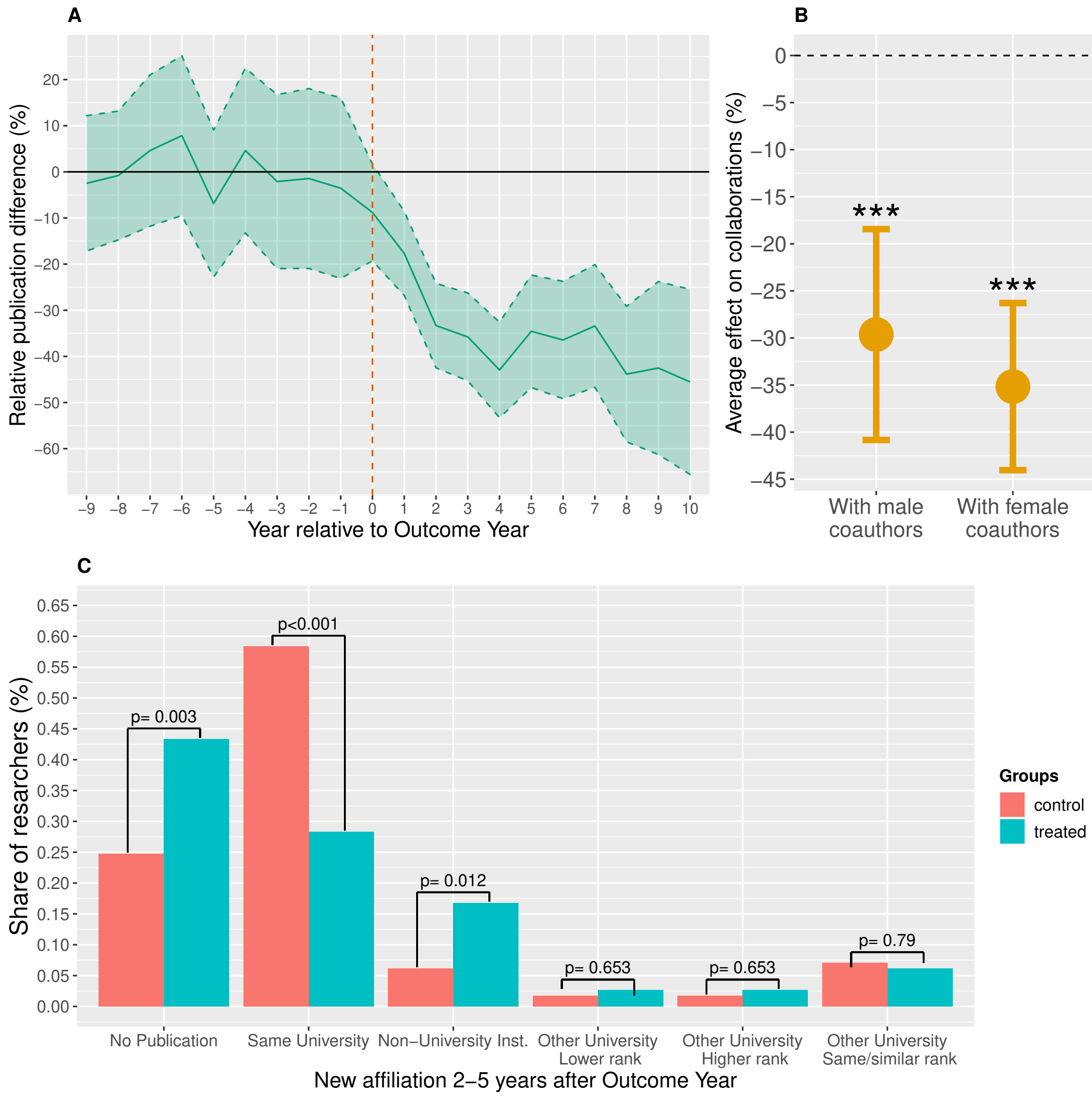
where $x_{s,t}$ is the number of published articles (number of other coauthors on published articles) (co)authored by scientist s in year t, transformed by the inverse hyperbolic sine, a_s is a scientist fixed effect, D_s denotes whether or not scientist s is accused and $I_{t,\tau^*(s)}$ is an indicator that captures the timing relative to the outcome year $\tau^*(s)$. We include a fixed effect for year γ_t and a fixed effect $\eta_{f(s,t)}$ for the academic age (years since the first publication) of scientist s (as of year t). We trim our data set ten years before and ten years after the outcome year. We estimate this model analogously to the article-level model in the previous section. Standard errors are clustered at the scientist level.

Figure 3 Panel A focuses on the number of publications. We find an average drop in yearly publications of -30.81% (t=-7.33, p<0.001) over (up to) 10 years after the outcome year.[14] The effect does not seem to weaken over time. Attrition from academic research appears to explain around four-fifths of the decline in publication output (accused scientists are 23.7% less likely to continue publishing, discussed below). For accused scientists who continue to publish, we do not find that the average quality of journals in which they publish declines.

The allegations might affect the accused scientist's ability to attract and retain coauthors, specifically when they are female. In Panel B, we report a decline in the number of collaborations with female researchers by -35.16% (t=-7.77, p<0.001). The point estimate of the effect on

[14] We limit the time horizon for effects on publications because researchers may retire.

Figure 3: Change in publication output and affiliation changes



Notes: Panel A shows decline in publication output per year, compared to the control group. Dotted lines indicate 95% confidence intervals. The average decline over (up to) 10 years after the outcome year is -30.81% (p<0.001). Panel B shows the decline in the number of coauthors of the specified gender on joint publications in the (up to) 10 years following the outcome year. Panel C shows the share of researchers that quit academic research ("no publication") and, for the remaining ones, the (latest available) affiliation 2-5 years after the outcome year. The new affiliation is compared to the affiliation held prior to the allegations.

collaborations with male researches is slightly smaller (in absolute terms), at around -29.63% (t=-5.19, p<0.001). The reduction in the number of both male and female coauthors can be explained by the reduction in the publication output of accused scientists.[15] Yet, the decrease in publication output may in turn be linked to difficulties in finding coauthors.

In Panel C, we consider affiliation changes, using information from articles published 2 to 5 years after the outcome year.[16] We find that accused scientists are around 18.6 percentage points more likely to leave science than control scientists (43.4% and 24.8% respectively), which we define as ceasing to publish.[17] This difference in attrition is statistically significant (p=0.003).

We classify the remaining scientists as either staying at the same university, moving to a non-university institution or moving to a lower-ranked/similar/higher-ranked university. We find that significantly fewer of the accused than control scientists stay at the same university (28.3% vs. 58.4%, p<0.001). Accused scientists exhibit a higher likelihood of publishing with a non-university affiliation (16.8% vs. 6.2%, p=0.012). We do not detect significant differences between the two groups in terms of moving to different universities, neither in the overall share (11.5% vs. 10.6%, p=0.833), nor disaggregated by the relative rank of old and new university.

4 Conclusion and Discussion

Previous studies have documented that, following a retraction, citation losses of about -6.9% to -10.8% spread to the researcher's entire body of prior work (Azoulay et al. 2017; Jin et al.

[15] When we restrict the sample to accused scientists who continue to publish, we do not find evidence of a change in the average number of female or male coauthors per published paper.

[16] During the initial 2 years after the event, it is conceivable that the coauthors may publish joint work from before the allegations, although the accused scientist may already be inactive. The presented results are for the sample of 113 accused scientists for whom affiliations are available and whose outcome year is before 2019. More details can be found in appendix C.3.

[17] The relative decline in the propensity to continue publishing obtains as $\left[1-(1-0.434)/(1-0.248)\right]*100=23.73\%$.

2019; Lu et al. 2013). We find that a similar citation penalty applies when incidents of sexual misconduct become public.[18] We estimate citation losses to the accused scientist's prior work of -5.25% due to the allegations.

We find that the intensity of the citing authors' response weakens with the distance in the coauthorship network. Our interpretation is that researchers learn about allegations through their peer network, and that common news sources (e.g. newspapers) play a negligible role in disseminating this information. Gender differences in citing behavior are limited to the closest peers, where, in line with sociological theories of "guilt by association" (Hussinger and Pellens 2019; Piazza and Perretti 2015; Pontikes et al. 2010), stronger responses by men may indicate that they feel a particular need to disassociate themselves from the accused scientist.

Prior literature suggests several other reasons why peers may choose to withhold citations. First, they may do so to penalize. It has been established that peers punish others to discourage antisocial behavior (e.g. Fehr and Gächter 2002; Henrich et al. 2006; Molho et al. 2020). Second, they may not cite the accused to avoid being seen as condoning sexual misconduct (Calvó-Armengol and Jackson 2010; Kandel and Lazear 1992). Third, they may suspect that sexual misconduct and scientific misconduct correlate, and therefore lose trust in the integrity of the work of the accused (similar to the findings of Griffin et al. 2019). Last, if they suspect that the accused abused his power to claim undue credit for his work with others (specifically with females, see Ross et al. (2022)), it may seem fair to stop citing him.

Our paper is the first to provide systematic evidence on the consequences of sexual misconduct for the accused. Our findings raise a number of ethical questions that highlight the tension between advancing knowledge and advancing science as a social institution. Is the decline in

[18] We find citation losses of -10.52%(p<0.001, t=-6.66) for the subsample of articles that appeared in journals in the top 10% of their field. This subsample is more comparable to the sample considered in studies of retractions, as retractions almost exclusively occur in top journals (Jin et al. 2019; Lu et al. 2013).

citations to the perpetrator's body of prior work an undue distortion of the scientific process or an appropriate penalty? Does society lose scientific output when it excludes or penalizes perpetrators? Are the documented career consequences adequate, also taking into account their possible deterrence effect? For example, should accused scientists be allowed to continue their career at places without teaching, such as non-university research institutions? Our results provide a new basis for a discussion of these important topics going forward.

References

Adams-Prassl, Abi, Kristiina Huttunen, Emily Nix, and Ning Zhang (2022). "Violence against women at work". In.

Antecol, Heather and Deborah Cobb-Clark (2006). "The sexual harassment of female active-duty personnel: Effects on job satisfaction and intentions to remain in the military". In: *Journal of Economic Behavior & Organization* 61.1, pp. 55–80.

Azoulay, Pierre, Alessandro Bonatti, and Joshua L Krieger (2017). "The career effects of scandal: Evidence from scientific retractions". In: *Research Policy* 46.9, pp. 1552–1569.

Baas, Jeroen, Michiel Schotten, Andrew Plume, Grégoire Côté, and Reza Karimi (Feb. 2020). "Scopus as a curated, high-quality bibliometric data source for academic research in quantitative science studies". In: *Quantitative Science Studies* 1.1, pp. 377–386. ISSN: 2641-3337. DOI: 10.1162/qss{_}a{_}00019. URL: https://www.mitpressjournals.org/doi/abs/10.1162/qss_a_00019.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess (2021). "Revisiting event study designs: Robust and efficient estimation". In: *arXiv e-print* (2108.12419).

BYU's Title IX Process (n.d.). https://titleix.byu.edu/byus-title-ix-process. Accessed: 2022-01-18.

Calvó-Armengol, Antoni and Matthew O Jackson (2010). “Peer pressure”. In: *Journal of the European Economic Association* 8.1, pp. 62–89.

Cantalupo, Nancy Chi and William Kidder (2018). “A systematic look at a serial problem: Sexual harassment of students by university faculty”. In: *Utah Law Review*, p. 671.

Cantor, David et al. (2019). *Report on the AAU campus climate survey on sexual assault and misconduct*. Report prepared for Westat. Rockville, Maryland: The Association of American Universities.

Chan, Darius KS, Suk Yee Chow, Chun Bun Lam, and Shu Fai Cheung (2008). “Examining the job-related, psychological, and physical outcomes of workplace sexual harassment: A meta-analytic review”. In: *Psychology of Women Quarterly* 32.4, pp. 362–376.

Cheng, Ing-Haw and Alice Hsiaw (2022). “Reporting sexual misconduct in the# MeToo era”. In: *American Economic Journal: Microeconomics* 14.4, pp. 761–803.

Cici, Gjergji, Mario Hendriock, Stefan Jaspersen, and Alexander Kempf (2021). “# MeToo meets the mutual fund industry: Productivity effects of sexual harassment”. In: *Finance Research Letters* 40, p. 101687.

Clancy, Kathryn BH, Lilia M Cortina, and Anna R Kirkland (2020). “Opinion: Use science to stop sexual harassment in higher education”. In: *Proceedings of the National Academy of Sciences* 117.37, pp. 22614–22618.

Dahl, Gordon B and Matthew M Knepper (Sept. 2021). *Why is Workplace Sexual Harassment Underreported? The Value of Outside Options Amid the Threat of Retaliation*. Working Paper 29248. National Bureau of Economic Research. DOI: 10.3386/w29248. URL: http://www.nber.org/papers/w29248.

Education Office for Civil Rights, US Department of (1997). “Sexual harassment guidance: Harassment of students by school employees, other students, or third parties”. In: *Federal Register* 62.49, pp. 12034–12051.

Egan, Mark, Gregor Matvos, and Amit Seru (2022). "When Harry fired Sally: The double standard in punishing misconduct". In: *Journal of Political Economy* 130.5, pp. 1184–1248.

Espinoza, Magdalena and David Hsiehchen (2020). "Characteristics of faculty accused of academic sexual misconduct in the biomedical and health sciences". In: *JAMA* 323.15, pp. 1503–1505.

Fehr, Ernst and Simon Gächter (2002). "Altruistic punishment in humans". In: *Nature* 415.6868, pp. 137–140.

Feldblum, Chai R and Victoria A Lipnic (2016). *Select task force on the study of harassment in the workplace*. Report of co-chairs. Washington, DC: U.S. Equal Employment Opportunity Commission. URL: `https://www.eeoc.gov/select-task-force-study-harassment-workplace`.

Fitzgerald, Louise F, Michele J Gelfand, and Fritz Drasgow (1995). "Measuring sexual harassment: Theoretical and psychometric advances". In: *Basic and Applied Social Psychology* 17.4, pp. 425–445.

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". In: *Journal of Vocational Behavior* 32.2, pp. 152–175.

Folke, Olle and Johanna Rickne (2022). "Sexual harassment and gender inequality in the labor market". In: *The Quarterly Journal of Economics* 137.4, pp. 2163–2212.

Fortney, Susan Saab and Theresa A Morris (2021). "Eyes Wide Shut: Using Accreditation Regulation to Address the "Pass-the-Harasser" Problem in Higher Education". In: *California Law Review Online*.

Gertsberg, Marina (2022). "The Unintended Consequences of# MeToo-Evidence from Research Collaborations". In: *Available at SSRN 4105976*.

González-Pereira, Borja, Vicente P. Guerrero-Bote, and Félix Moya-Anegón (2010). “A new approach to the metric of journals scientific prestige: The SJR indicator”. In: *Journal of Informetrics* 4.3, pp. 379–391. ISSN: 17511577. DOI: 10.1016/j.joi.2010.03.002.

Griffin, John M, Samuel Kruger, and Gonzalo Maturana (2019). “Personal infidelity and professional conduct in 4 settings”. In: *Proceedings of the National Academy of Sciences* 116.33, pp. 16268–16273.

Henrich, Joseph et al. (2006). “Costly punishment across human societies”. In: *Science* 312.5781, pp. 1767–1770.

Hussinger, Katrin and Maikel Pellens (2019). “Guilt by association: How scientific misconduct harms prior collaborators”. In: *Research policy* 48.2, pp. 516–530.

Jin, Ginger Zhe, Benjamin Jones, Susan Feng Lu, and Brian Uzzi (2019). “The reverse Matthew effect: Consequences of retraction in scientific teams”. In: *Review of Economics and Statistics* 101.3, pp. 492–506.

Johnson, Laura S (2015). “Gender discrimination and Title IX implementation: Lessons from the Office for Civil Rights resolution letters 1997-2011”. PhD thesis. University of Kentucky. URL: https://uknowledge.uky.edu/cgi/viewcontent.cgi?article=1034&context=epe_etds.

Kabat-Farr, Dana and Lilia M Cortina (2014). “Sex-based harassment in employment: New insights into gender and context.” In: *Law and Human Behavior* 38.1, p. 58.

Kandel, Eugene and Edward P Lazear (1992). “Peer pressure and partnerships”. In: *Journal of Political Economy* 100.4, pp. 801–817.

Koebel, James T (2016). “Campus misconduct proceeding outcome notifications: A Title IX, Clery Act, and FERPA compliance blueprint”. In: *Pace L. Rev.* 37, p. 551.

Koffi, Marlène (2021). “Gendered citations at top economic journals”. In: *AEA Papers and Proceedings*. Vol. 111, pp. 60–64.

Kotkin, Minna J (2005). “Invisible Settlements, Invisible Discrimination”. In: *North Caroline Law Review* 84, p. 927.

Laband, David N and Bernard F Lentz (1998). “The effects of sexual harassment on job satisfaction, earnings, and turnover among female lawyers”. In: *ILR Review* 51.4, pp. 594–607.

Libarkin, J. (2019). *Academic Sexual Misconduct Database*. URL: https://academic-sexual-misconduct-database.org (visited on 10/10/2019).

Lisak, David, Lori Gardinier, Sarah C Nicksa, and Ashley M Cote (2010). “False allegations of sexual assault: An analysis of ten years of reported cases”. In: *Violence against women* 16.12, pp. 1318–1334.

Lu, Susan Feng, Ginger Zhe Jin, Brian Uzzi, and Benjamin Jones (2013). “The retraction penalty: Evidence from the Web of Science”. In: *Scientific Reports* 3.1, pp. 1–5.

Lucey, Jemi G., Cameryn J. Hinton, and Irene Hsieh (2020). *Sexual Violence and the New Title IX Rules: Where Do We Go from Here?* Report. URL: https://www.greenbaumlaw.com/media/publication/715_NJLJ-Reprint-TitleIX-JGL-IH-CJH.pdf.

Luo, Hong and Laurina Zhang (2022). “Scandal, social movement, and change: Evidence from# metoo in hollywood”. In: *Management Science* 68.2, pp. 1278–1296.

Maestas, Nicole, Kathleen J Mullen, David Powell, Till Von Wachter, and Jeffrey B Wenger (2017). “Working Conditions in the United States: Results of the 2015 American Working Conditions Survey”. In: *RAND Corporation*.

McCann, Carly, Donald Tomaskovic-Devey, and M.V. Lee Badgett (2018). “Employer’s Responses to Sexual Harassment”. In: *Available at SSRN 3407960*.

McLaughlin, Heather, Christopher Uggen, and Amy Blackstone (2017). “The economic and career effects of sexual harassment on working women”. In: *Gender & Society* 31.3, pp. 333–358.

Melnick, R. Shep (June 11, 2020). *Analyzing the Department of Education’s final Title IX rules on sexual misconduct*. Brookings. URL: https://www.brookings.edu/research/

analyzing-the-department-of-educations-final-title-ix-rules-on-sexual-misconduct/ (visited on 01/18/2022).

Merton, Robert K. (1973). *The Sociology of Science: Theoretical and Empirical Investigations*. Chicago: University of Chicago Press.

Moed, Henk F., M'hamed Aisati, and Andrew Plume (Mar. 2013). "Studying scientific migration in Scopus". In: *Scientometrics* 94.3, pp. 929–942. ISSN: 0138-9130. DOI: 10.1007/s11192-012-0783-9. URL: http://link.springer.com/10.1007/s11192-012-0783-9.

Molho, Catherine, Joshua M Tybur, Paul AM Van Lange, and Daniel Balliet (2020). "Direct and indirect punishment of norm violations in daily life". In: *Nature Communications* 11.1, pp. 1–9.

National Academies of Sciences, Engineering, and Medicine (2018). *Sexual Harassment of Women: Climate, Culture, and Consequences in Academic Sciences, Engineering, and Medicine*. Washington, DC: The National Academies Press. ISBN: 978-0-309-47087-2. DOI: 10.17226/24994. URL: https://www.nap.edu/catalog/24994/sexual-harassment-of-women-climate-culture-and-consequences-in-academic.

Piazza, Alessandro and Fabrizio Perretti (2015). "Categorical stigma and firm disengagement: Nuclear power generation in the United States, 1970–2000". In: *Organization Science* 26.3, pp. 724–742.

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". In: *American sociological review* 75.3, pp. 456–478.

Rose, Michael and Stefano Baruffaldi (2020). "Finding Doppelgängers in Scopus: How to Build Scientists Control Groups Using Sosia". In: *Max Planck Institute for Innovation & Competition Research Paper* 20. ISSN: 1556-5068. DOI: 10.2139/ssrn.3742602. URL: https://www.ssrn.com/abstract=3742602.

Rose, Michael E. and John R. Kitchin (July 2019). "pybliometrics: Scriptable bibliometrics using a Python interface to Scopus". In: *SoftwareX* 10.July-December, p. 100263. ISSN: 23527110. DOI: 10.1016/j.softx.2019.100263. URL: https://linkinghub.elsevier.com/retrieve/pii/S2352711019300573.

Ross, Matthew B, Britta M Glennon, Raviv Murciano-Goroff, Enrico G Berkes, Bruce A Weinberg, and Julia I Lane (2022). "Women are credited less in science than men". In: *Nature* 608.7921, pp. 135–145.

Sample Language for Reporting and Confidentially Disclosing Sexual Violence (n.d.). https://www.justice.gov/file/910281/download. Accessed: 2022-01-18.

Sarsons, Heather, Klarita Gërxhani, Ernesto Reuben, and Arthur Schram (2021). "Gender differences in recognition for group work". In: *Journal of Political Economy* 129.1, pp. 101–147.

A Additional Information on Sexual Misconduct Allegations

A.1 Legal background

The legal environment surrounding sexual misconduct at U.S. universities affects the reporting of allegations, which allegations are investigated and how they are handled by the university. The treatment of sexual misconduct cases in educational institutions is governed by Title IX of the Education Amendments of 1972. Title IX is an equity law that focuses on issues of gender discrimination. It was not specifically developed to address sexual misconduct, but starting from mid 1990s its applicability expanded to include it (Johnson 2015). Educational institutions faced increased pressure to show greater attention to curbing sexual misconduct on their campuses. To help educational institutions to implement policies that comply with Title IX, the Office for Civil Rights (OCR) issued a series of letters. The first such letter – "Sexual Harassment Guidance: Harassment of Students by School Employees, Other Students, or Third Parties"– was issued in 1997. Effectively, it was the first federal guidance on the implementation of Title IX policies in application to sexual harassment and all future guidance builds on this one (Johnson 2015; Lucey et al. 2020). We allow for a time lag of one year for new processes to be implemented and consider the cases starting from 1998.

All public universities in the U.S., as well as those that receive grants from NSF, NIH or NASA, must have Title IX policies in place that, among others, inform students and staff about issues of gender equality and introduce ways of reporting incidents of sexual misconduct. Universities are required to react to all reports of sexual misconduct. However, not all reports trigger formal investigations (*BYU's Title IX Process* n.d.; Melnick 2020; *Sample Language for Reporting and*

Confidentially Disclosing Sexual Violence n.d.):[19] The responsible office decides whether a formal investigation is compulsory based on the seriousness of the allegations, the age or maturity of the complainant, the existence of any previous accusations against the alleged perpetrator, and the existence of independent evidence to substantiate the allegations. In all other cases, there is a possibility that parties are offered to resolve the complaint informally, which leaves no formal finding and does not result in disciplinary action by the university.

We do not have researchers in our sample whose names are suppressed by confidentiality or privacy regulations. These regulations can operate at different institutional levels (state, university, bilateral contracts) (Cantalupo and Kidder 2018; Koebel 2016; Kotkin 2005). However, to the best of our knowledge, there is no reason to believe that the availability of names is systematically related to the characteristics of the accused.

A.2 Substantiation of allegations and guilt of the accused

We believe that our sample focuses on cases that appeared substantiated. As described in section A.1, the availability of corroborating evidence is part of the criteria for launching an investigation. The inclusion criteria of our data sources, as we discuss in section 2, further indicate that the allegations must have been credible. Yet, at the same time, not all criteria unambiguously signify a finding of guilt or an admission of guilt.

Prior research suggests that false reporting of sexual harassment is rare. McCann et al. (2018) show that 88% of reports of sexual harassment discrimination charges submitted to the U.S. Equal Employment Opportunity Commission "have legal merit". Lisak et al. (2010) documents rates of false allegations of sexual assault to be between 2% and 10% based on studies in differ-

[19] The procedure and requirements for formal investigation changed in 2020 following the adoption of Education Department's Final Title IX Rule. This change, however, happened outside of the time period when it could have affected our sample.

ent countries and time periods. In an educational setting, Lisak et al. (2010) find a rate of false reports of 5.9% in sample of reports of sexual assault submitted to the police department of a major university.

We conduct a supplementary review to estimate the share of cases in our sample where the allegations were found false. To this end, we form a random sample of 40 cases and read all the coverage in LexisNexis. We searched for statements on the outcome of investigation and if the allegations themselves were upheld. For 60% of the sampled cases (24 total), we are able to find a definite statement of guilt made by the university or a court (e.g. "X's conduct violated university policies on non-discrimination, anti-harassment, non-retaliation, sexual harassment and the university's code of ethics" or "a university investigation found him guilty of sexual harassment"). For 30% (12 cases), we a) do not find a definite statement of guilt, but b) we also do not find a definite statement to the contrary, i.e. no acquittal of the accused. In four cases (10%), the university investigation acquitted the accused (e.g. "The investigation found the professor had not violated the school's sexual harassment policy."). However, even then, this does not mean that the allegations were false, but that they did not fulfill the legal requirements to qualify as sexual misconduct. For example, in one of the four cases, the accused was found to have acted unethically, albeit not illegally, and fired from his position. In none of the cases that we reviewed the allegations themselves were proven false.

We therefore estimate that the share of cases in our sample in which the allegations qualify as sexual misconduct and are unambiguously true as 60-90%. False allegations per se constitute, if at all, a negligible share.

A.3 Classification of largest U.S. newspapers

An incident may gain additional visibility and thus become known to more people if it is covered by a newspaper with large readership. We proxy readership by the circulation numbers of newspapers and classify 33 U.S. newspapers with highest circulation numbers as "large". Additionally, we classify incidents that were covered by Associated Press, one of the U.S. largest news organizations, as the ones with broad readership.

33 largest U.S. newspapers by circulation represent a combined list of the top 25 newspapers in the U.S. for the year 2007 by Audit Bureau Circulation and for the year 2021 by Alliance for Audited Media. It includes Wall Street Journal, the New York Times, USA Today, Washington Post, New York Post, Los Angeles Times, Chicago Tribune, Star Tribune, Tampa Bay Times, Newsday, Seattle Times, Honolulu Star-Advertiser, Arizona Republic, Boston Globe, Dallas Morning News, Houston Chronicle, Philadelphia Inquirer, San Francisco Chronicle, Denver Post, Chicago Sun-Times, The Buffalo News, Daily News, Villages Daily Sun, St. Louis Post-Dispatch, Milwaukee Journal Sentinel, Newark Star-Ledger, Atlanta Journal-Constitution, Plain Dealer, News/Free Press Detroit, Oregonian (Portland), Union Tribune San Diego, the California Register, Sacramento Bee.

B Bibliometric Data

B.1 Matching with the Scopus database

All our bibliometric data originates from Elsevier's Scopus database. Three of its features are important for the current analysis. First, Scopus maintains disambiguated profiles with unique identifiers for researchers and institutions of high quality.[20] Disambiguated author profiles are important for the construction of co-author networks and for assigning complete publication records, while disambiguated institutions allow us to trace researcher's careers. Second, Scopus' coverage is not restricted to certain fields or disciplines. Third, Scopus assures integrity of data and, in combination with the code of Rose and Kitchin (2019), replicability of results by providing a RESTful API.

We match accused scientists to their author profiles on Scopus based on name, affiliation and field of research. Whenever the publications of accused scientists are split across multiple profiles, we aggregate profiles. We carry out complementary checks in indeterminate cases to avoid false matches.

B.2 Publications data

For each accused scientist, we retrieve the entire list of research-type publications and exclude books and book chapters.[21] For each publication, we obtain the journal name, the publication year, and the listed affiliations and names of the authors. We track all subsequent citations (excluding self-citations) to the publications (co-)authored by accused scientists as of January 2021 and record the authors of citing articles.

Research fields We assign publications to research fields using Scopus' "All Science Journal Classification" (ASJC). The ASJC defines 26 research fields (2-digit level) with "Multidisci-

[20] Baas et al. (2020) estimate the precision of Scopus author profiles equal to 98.1% and recall equal to 94.4%. The algorithm thus favors excess profiles over merging too many publications into a false profile (Moed et al. 2013).

[21] We exclude books and book chapters because they are incompatible with our control group matching strategy (see Appendix B.4)

plinary" being a separate field.[22] If a journal is assigned to more than one research field, we pick one at random. However, we deviate from this rule when we classify journals as belonging to male-dominated/not-male-dominated fields (see the paragraph Gender below).

Gender We obtain a (probabilistic) gender estimate for all researchers in Scopus, including the accused scientists, all of their co-authors and all citing authors from the service gender-api.com.[23] We use this information to establish the gender distribution of co-authors and citing authors.

Furthermore, we compute the overall author gender-ratio for each research field by year. We define a field as male-dominated if it has a share of female authors of below 20% in the year 2000. This includes Engineering, Energy, Physics and Astronomy, Computer Science, Mathematics, Economics, Decision Science, Earth and Planetary Science, Materials Science and Chemical Engineering.[24] The relative ranking of fields in terms of author gender-ratio share appears stable over time. We assign an article to the male-dominated fields if weakly more than 50% of the research fields that are assigned to the journal are male-dominated (in case the journal belongs to multiple research fields). A more restrictive definition for defining a male-dominated field and for assigning articles to male-dominated fields is considered in Appendix C.2.

Journal ranking We assign journals a percentile-rank within their research field based on the SCImago Journal ranking (SJR).[25] This "size-independent indicator of journals' scientific pres-

[22]See `https://service.elsevier.com/app/answers/detail/a_id/15181/supporthub/scopus/` for a complete list of all codes, fields and subject areas.

[23]The estimation is based on the first names. Currently 6 million validated names are present in their database which originate mainly from public records. See `https://gender-api.com/`.

[24]We do not calculate gender-ratios for the field "Multidisciplinary" (which contains general-interest journals such as Science, Nature or PNAS).

[25]SJR indicators are based on Scopus data and can be found at `http://www.scimagojr.com/journalrank.php`.

tige" is an iterative Eigenscore that weights citations to individual documents of the past three years by the SJR indicator of the citing articles (González-Pereira et al. 2010). We interpolate the SJR indicator for years in which it is missing.

Affiliations We use affiliations recorded on publications to construct the affiliation history of researchers. We aggregate intermediate units (e.g. hospitals, libraries, departments, etc.) at the institution level and categorize by type (university vs non-university). We use this data to detect affiliation changes and transitions between the university and the non-university sector (described in Appendix C.2). In addition, we match universities by university name to the Shanghai university ranking.

B.3 Coauthorship network

For each pair of citing and focal paper, we determine the minimum distance in the global coauthorship network between the authors listed on the two papers. More precisely, let x be the focal article with a set of authors $\mathscr{A}(x)$, and let c be the citing article with the set of authors $\mathscr{A}(c)$. If x is the article of an accused, the relevant author j is the accused; otherwise it is a random author chosen from $\mathscr{A}(x)$. Then the distance is given by

$$d_{xc} = \min \tilde{d}_{ij} \ \forall i \in \mathscr{A}(c). \tag{3}$$

where $\tilde{d}_{ij}$ is the geodesic distance between two authors in the network. If there is no path between two authors, they are said to be in distinct components and $\tilde{d}_{ij}$ is set to $+\infty$ (which would then imply that d_{xc} takes on the value $+\infty$ too).

The relevant coauthorship network for each pair of (x,c) depends on the publication year of c

and the outcome year of the allegation. We construct one network for each year $\in (1971, 2020)$, which includes all journal publications in Scopus published between 1947 and the year before the publication of c or the outcome year, whichever comes first. By fixing the coauthorship network prior to the outcome year, we abstract from endogenous adjustments in the network due to the allegations and evaluate subsequent citations relative to the network that was present when allegations surfaced.

In principle, we assume that two authors are connected in a given year if they had ever published an article as coauthors prior to this point. However, we remove "inactive" researchers from the respective network, i.e. those that stopped publishing for at least two years prior to the focal year. Our earliest network in 1971 includes around 0.9 million authors connected by around 2.2 million links, while our latest network in 2019 contains 10.6 million authors connected by 158 million links.

B.4 Matching: Control articles and control scientists

Control articles In the first part of our analysis, we compare the citations received by a set of articles (co-)authored by accused scientists (referred to as "treated" articles) to a control group drawn from the same journal issue (referred to as control articles). If the periodical does not use issues, we sample from the entire volume. The control group consists of up to 10 randomly selected research articles published in the same issue (provided their authors are not accused scientists too). If multiple accused scientists publish in the same issue/volume, we draw control groups independently. We track subsequent citations to control articles in the same way as for treated articles and determine the gender (see Section B.2) and coauthorship distance of citing authors (see Section B.3).

Control Scientists In the second part of the analysis, we compare accused scientists to other observationally similar scientists. To identify them, we use the Python package "sosia" by Rose and Baruffaldi (2020). "sosia" finds academic "twins" based on a number of observable characteristics in the Scopus database. Treated and control scientists are matched based on characteristics in the year equal to $\max(t_{\text{outcome}} - 4, t_{\text{incident}} - 1)$, which precedes the outcome year by between one and four years. The choice of this time lead is motivated by the concern that a match based on a later year might be influenced by changes in performance due to allegations.

We match the accused scientist to a control scientist who:

1. publishes in the sources (journals or proceedings) belonging to main research areas of the accused scientist,
2. started publishing in the same year as the accused scientist plus/minus three years,
3. has the same stock of publications as of the matching year within a margin of 30%,
4. has the same stock of distinct coauthors as of the matching year within a margin of 30%,
5. has received the same number of citations overall (plus/minus 500),
6. whose main affiliation (most commonly used affiliation during the matching year) is located in the same country,
7. is at an institution of the same type (university, research institute, independent hospital, governmental organization, non-profit organization, industry) and
8. is of the same gender.

The procedure yields matches for 158 accused scientists (i.e., 54 without any match). Most of them (142) have multiple viable matches, up to 271. If several matches are suggested, we choose one randomly.

C Methods

C.1 Additional information

Estimation details In our analysis, we use the estimation strategy proposed by Borusyak et al. (2021), which is robust and efficient when the treatment effect depends on the time of treatment. We implement the estimator using the STATA-package *did_imputation*. We estimate the unweighted average of all time- and article-specific treatment coefficients after the outcome year $\bar{\tau} = \frac{1}{N}\sum_{i,t}\{\tau_{i,t}|t \geq \tau^*(i)\}$ and the averages of the treatment coefficients for each year relative to the outcome year, given by $\bar{\tau}_T = \frac{1}{N_T}\sum_{i,t}\{\tau_{i,t}|t = \tau^*(i) + T\}$. The estimates are converted to percent semi-elasticities using the transformation $\left(\sinh(\sinh^{-1}\bar{c} + \hat{\tau})/\bar{c} - 1\right) * 100$, where $\bar{c}$ is the mean of the annual citation counts in the sample (annual publication or collaboration counts in the scientist-level analysis). The corresponding standard errors are obtained with the Delta method and clustered at the level of the accused scientist (in the treatment group) or at the article level (control group).

Tables S5, S6 and S7 show the results underlying Figure 2. We also present estimates from a standard Poisson panel model where all treatment coefficients are assumed to be homogeneous, using the package *ppmlhdfe*. These estimates are converted into percent semi-elasticities using the transformation $\left(\exp(\hat{\tau}) - 1\right) * 100$. All standard errors are obtained using the Delta method.

Citation trend graphs In Figure S4, we present citation trend graphs for articles by accused scientists and control articles. In the top panel, we use all incidents; however, this means that starting three years after the outcome year, we reach the truncation point for some incidents (the most recent cases have an outcome year of 2019, and we track citations through 2021), which are then dropped for the later years. In the bottom panel, we restrict the sample to incidents with an outcome year before 2017, and track citations only for the first five years after the outcome year.

C.2 Robustness for results on male-dominated fields

We first alter the definition of "male-dominated" fields to include only those are that in the bottom quartile in terms of female author share (Engineering, Energy, Physics and Astronomy, Computer Science, Mathematics and Economics). This definition is more restrictive. Second, we make our assignment of articles to male-dominated fields more restrictive and require that a strict majority (instead of a weak majority) of $> 50\%$ of research fields that are assigned to the journal are male-dominated.

When we use the more restrictive definition of male-dominated fields in conjunction with the more restrictive journal assignment, we find an estimated effect of +4.11% (t=2.39, p=0.017) for articles that are published in male-dominated fields, and an effect of -5.32% for the remaining articles (t=-5.57, p<0.001). Using the more restrictive definition of male-dominated fields, but keeping the journal assignment as in our main specification, we obtain estimates of -2.94% (t=-1.00, p=0.319) and -5.12% (t=-5.59, p<0.001) for articles in male-dominated and not-male-dominated fields respectively. Last, using the more restrictive journal assignment, while keeping the original definition of male-dominated fields, yields estimates of -1.16% (t=-0.38, p=0.707) for male-dominated fields and -5.39% (t=-6.06, p<0.001) for not-male-dominated fields.

C.3 Analysis of scientist affiliation changes

First, we ascertain the affiliation for each accused scientist and his matched control scientist prior to the incident. We use the affiliation recorded in the base year of our matching procedure (see B.4 for details). If not available, we instead use the latest recorded affiliation from all publications that appeared in the 5 years preceding this year. Using this procedure, we obtain affiliation information for 131 pairs of treated and matched control scientist (out of 158).

Second, for each pair of treated and matched control scientist, we determine the (latest available) affiliations 2 to 5 years after the outcome year. Because our sample ends in 2021, we drop 18 pairs with an outcome year after 2018. Provided that there is at least one publication during this period, we classify the affiliation as either being identical to the pre-allegation affiliation, as belonging to a different university, or as belonging to a non-university institution. If the affiliation data indicates a transition between two different universities, we determine the difference in the Shanghai university ranking. Whenever the rank (or the rank-range)[26] of the new university is smaller (larger) than the rank (or the rank-range) of the former university, we classify the transition as a move to a "higher ranked university" (move to a "lower ranked university"). We classify transitions between universities of the same rank-range as moves between universities of "same/similar rank". The p-values that are reported in Figure 3 Panel C are based on the test $H_0 : b = 0$ in the regression $\tilde{x}_s = a + bD_s + e_s$, where $\tilde{x}_s$ is the indicated outcome of scientist s and D_s is the treatment indicator.

[26]We use the Top 500 of the Shanghai University ranking from 2004 to 2020. Universities outside of the top 100 do no receive individual ranks in the Shanghai ranking, but a rank-range. Universities that are not present in the ranking are classified as "unranked", which corresponds to the lowest rank-range. We compare university ranks when the new affiliation is recorded. Rank comparisons for years prior to 2004 are based on the 2004-ranking.

D Additional Tables and Figures

Table S1: Search terms for Academic Sexual Misconduct database (reproduced)

	search terms
Academic modifiers (one of)	'professor', 'instructor' AND 'university OR college', 'dean','university OR college' AND 'president', 'university OR college' AND 'provost', 'university OR college' AND 'administrator', 'university OR college' AND 'employee'
Combined via AND with	
additional modifer (one of)	[name of each US state], 'community college', 'tribal college'
AND	
sexual misconduct modifier (one of)	'sexual harassment', 'sexual misconduct', ,'sexual assault','inappropriate relationship', 'peeping','voyeur','rape','kidnap','murder'

Table S2: The nature of the incidents.

	Count	Share (in %)
Inappr. Comments	29	13.7
Unwanted Sex. Advances incl. touching	103	48.6
Stalking (incl. cyber)	6	2.8
Sex. relations	48	22.6
Rape	4	1.9
Other	13	6.1
Unknown	9	4.2

Notes: Under *inappropriate comments* we classified cases were allegations referred to verbal abuse without physical contact. *Unwanted sexual advances including inappropriate touching*: The allegations involve inappropriate touching, sending of sex related messages and offering sex but none of the complainants alleged having an intercourse with the accused. Incidents were coded as *rape* only if sources used this term.

Table S3: The outcome of the incidents.

	Count	Share (in %)
No consequences	2	0.9
Mild administrative measures	7	3.3
Monetary consequences	9	4.2
Serious administrative measures	40	18.9
Left the position	128	60.4
Criminal plea/conviction	13	6.1
Suicide	5	2.4
Other	3	1.4
Unknown	5	2.4

Notes: As *mild administrative* consequences we classified outcomes as, for instance, an official warning, reprimation, mandated training or removed honorable positions. *Monetary consequences* include paying settlement or fine, facing salary reduction. *Serious administrative measures* include suspension, probation, teaching restrictions, demotion. *"Left the position"* includes firing or resigning, contract non-renewal or retirement.

Table S4: Comparison of treated scientists with their matched control scientists.

	Treated scientists		Control scientists		Difference		
	Mean	Std. dev.	Mean	Std. dev.	Diff.	t-statistic	p-values
Year of First Publication	1992.29	11.33	1992.04	11.47	0.25	0.19	0.85
Career Age at Treatment	19.43	10.80	19.68	10.97	-0.25	-0.20	0.84
Publications	54.22	78.12	52.53	76.31	1.70	0.20	0.85
Citations	1224.37	2811.94	1203.32	2844.09	21.05	0.07	0.95
Coauthors	94.46	165.66	101.18	198.32	-6.72	-0.33	0.74

Table S5: Estimates of the effect of sexual misconduct allegations on citations by gender of the first author of the citing article

Citations by gender	All	Male	Female
Model A: Borusyak et al., DV: $\sinh^{-1}$(cit count)			
treatment_post (%)	-5.246	-6.013	-3.948
	(0.951)	(0.976)	(1.057)
Model B: Poisson, DV: cit count			
treatment_post (%)	-9.607	-10.327	-8.048
	(2.215)	(2.269)	(2.290)
Article Fixed Effects	✓	✓	✓
Journal Issue× Citation Year Fixed Effects	✓	✓	✓
Observations	2,271,982	2,271,982	2,271,982
Articles	108,471	108,471	108,471
Incidents	212	212	212

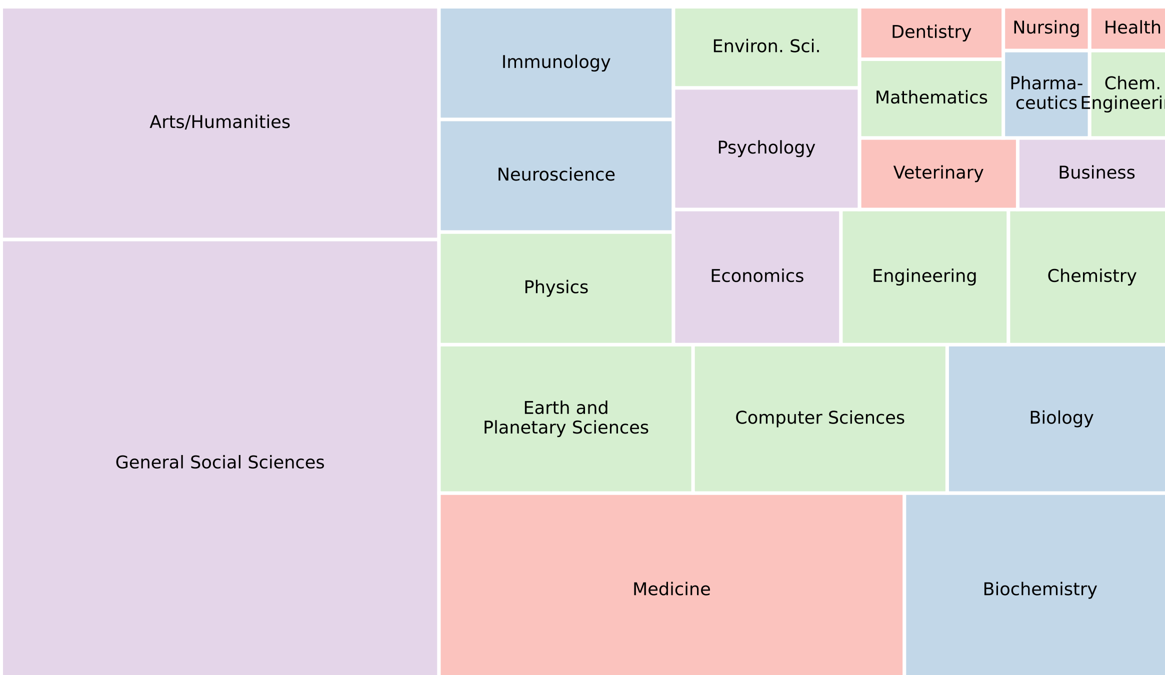
Table S6: Estimates of the effect of sexual misconduct allegations on citations by coauthorship distance and gender

Cit. by coauthorship dist.	Dist=1			Dist=2			Dist≥3		
Citations by gender	All	Male	Female	All	Male	Female	All	Male	Female
Model A: Borusyak et al., DV: $\sinh^{-1}$(cit count)									
treatment_post (%)	-28.743	-38.529	-19.800	-7.985	-10.997	-8.634	-0.717	-0.568	-0.955
	(3.967)	(5.067)	(5.648)	(1.534)	(1.840)	(2.502)	(1.939)	(1.446)	(1.374)
Model B: Poisson, DV: cit count									
treatment_post (%)	-21.573	-23.356	-15.886	-7.799	-8.314	-7.003	-3.372	-3.333	-3.863
	(5.043)	(5.212)	(5.787)	(3.107)	(3.126)	(3.404)	(2.329)	(2.504)	(2.173)
Article Fixed Effects	✓	✓	✓	✓	✓	✓	✓	✓	✓
Journal Issue× Citation Year Fixed Effects	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations		2,271,982			2,271,982			2,271,982	
Articles		108,471			108,471			108,471	
Incidents		212			212			212	

Table S7: Estimates of the effect of sexual misconduct allegations on citations by field of publication

Articles by field of publication	Not male-dominated fields			Male-dominated fields		
Citations by gender	All	Male	Female	All	Male	Female
Model A: Borusyak et al., DV: $\sinh^{-1}$(cit count)						
treatment_post (%)	-5.733	-6.524	-4.355	-0.813	-1.318	1.092
	(0.931)	(0.963)	(1.026)	(2.349)	(2.131)	(3.215)
Model B: Poisson, DV: cit count						
treatment_post (%)	-8.890	-9.480	-7.457	-7.086	-7.568	-7.845
	(2.200)	(2.263)	(2.332)	(4.822)	(4.806)	(5.363)
Article Fixed Effects	✓	✓	✓	✓	✓	✓
Journal Issue × Citation Year Fixed Effects	✓	✓	✓	✓	✓	✓
Observations		1,698,874			450,365	
Articles		80,124			22,482	
Incidents		209			107	

Figure S1: Distribution of incidents over scientific fields



Notes: Primary fields of perpetrators (N=212), as given by the most frequent 2-digit ASJC-code (All Science Journal Classification) of journals where their publications got published. Fields are color-coded by main area (Social Sciences, Physical Sciences, Life Sciences, Health Sciences).

Figure S2: Distribution of outcome years

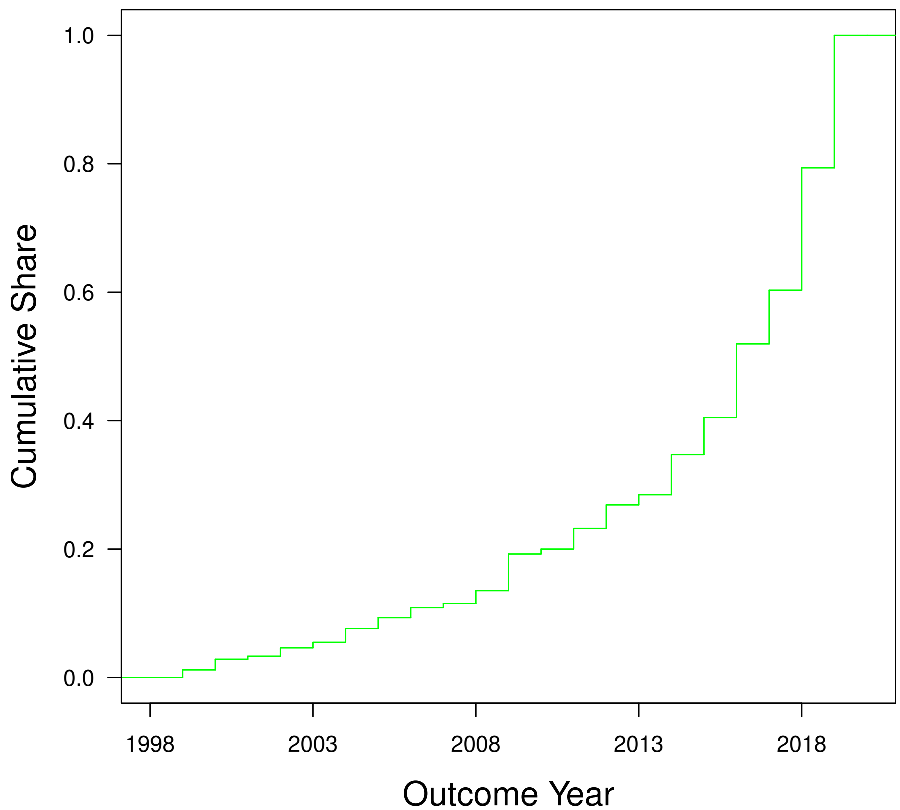
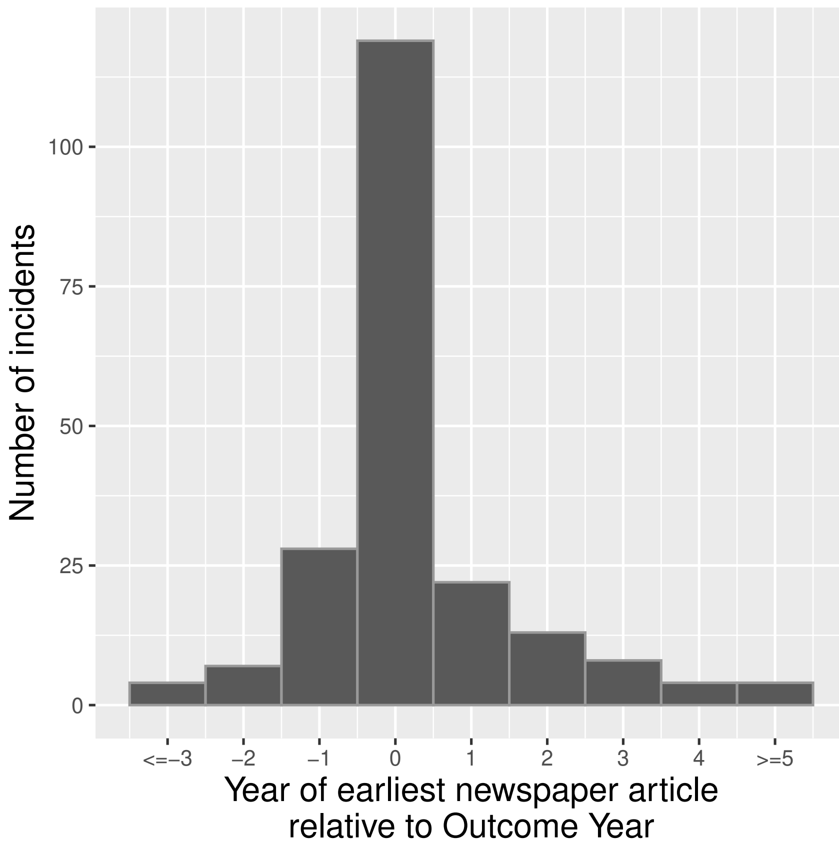
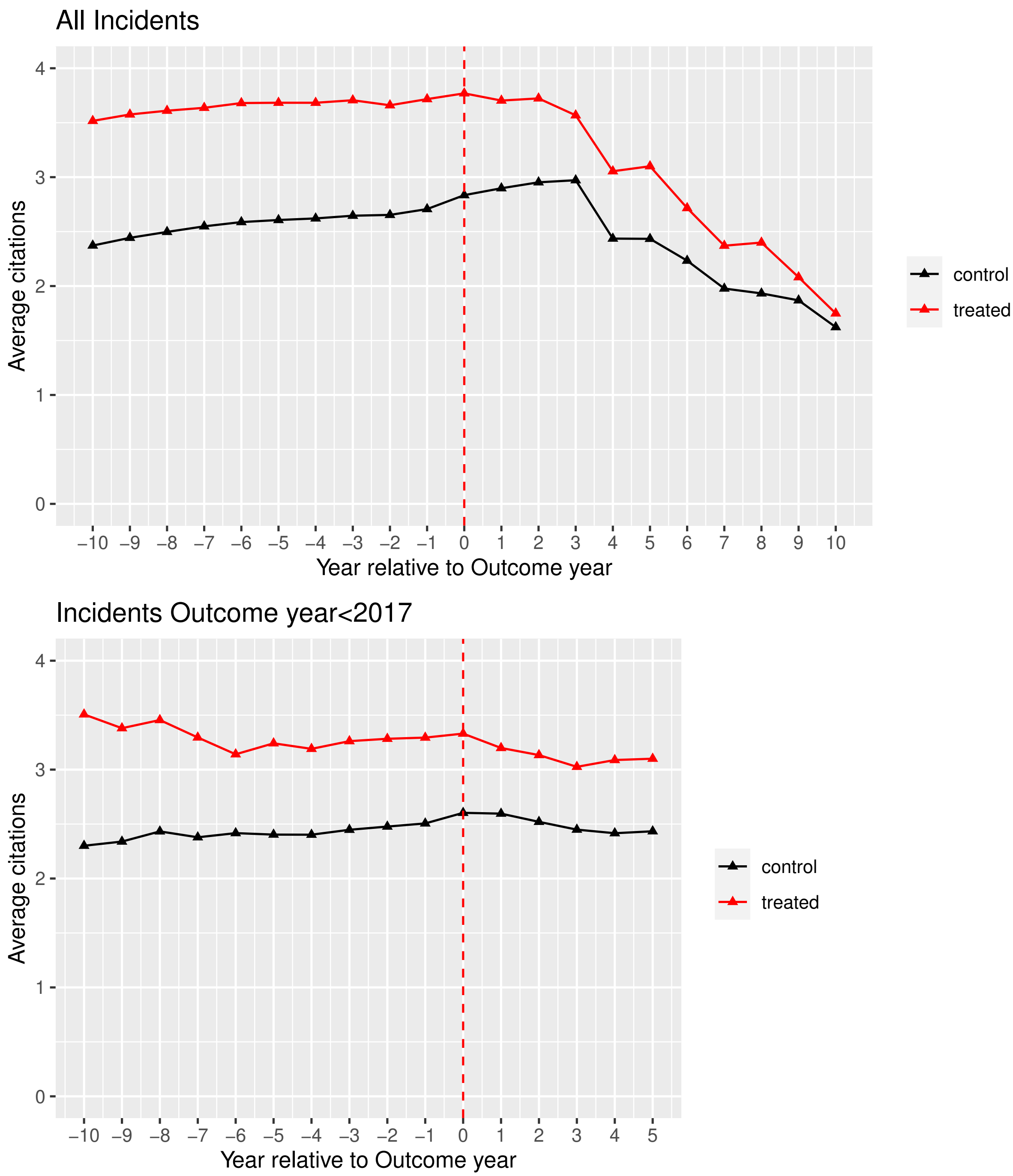


Figure S3: Distribution of time difference between publication of earliest newspaper article covering an incident and the outcome year



Notes: Negative time difference indicates that the earliest newspaper article predates the outcome year. 209 out of 212 incidents are covered by at least one newspaper article.

Figure S4: Citation trend graphs for treated and control articles



Notes: Average number of citations per year for articles of accused scientists published before the outcome year and control articles.