Cancel Culture: The Professional Costs of Speech Sanctioning Events Evidence from Academia*

Kris Gulati[a,b] Lorenzo Palladini[c]

[a] *Berkeley-Haas*

[b] *University of California, Merced*

[c] *Stockholm School of Economics*

January 20, 2026

ABSTRACT

Public controversies over employee speech where individuals face backlash for expressing views deemed controversial, unpopular, or misaligned with prevailing norms - sometimes colloquially referred to as "cancel culture" - have become increasingly prevalent. Yet, systematic evidence of their professional consequences remains limited. Leveraging a novel dataset of U.S.-based academics involved in such events, combined with bibliometric data, we employ a series of difference-in-differences designs to measure the effects on individuals' productivity and recognition. We find that, following these events, affected scholars experience a decline in the recognition of their prior body of work, with citations declining by 4.8%, reflecting a form of peer-to-peer distancing. Furthermore, affected scholars experience a decline in their productivity, publishing 20% fewer papers compared to a comparable set of scholars. These citation declines are disproportionately driven by scholars who are closely connected to the affected individual, consistent with a mechanism of professional distancing. Our findings highlight the professional and career costs associated with speech sanctioning events.

*Kris Gulati thanks the following organisations for their financial support: the Institute of Humane Studies under grant numbers: IHS018498, IHS018315, IHS018545; Emergent Ventures; and the Foundation for Individual Rights and Expression (FIRE). We are deeply grateful to Pierre Azoulay, Benjamin Balseimer, Tyler Cowen, Mathijs De Vaan, Solène Delecourt, Lee Fleming, Christian Fons-Rosen, Komi Frey, Matt Gentzkow, Paula Giuliana, Nate Honeycutt, Katrin Hussinger, John Ioannidis, Andrew Johnston, Ben Jones, Karim Lakhani, Saqib Mumtaz, Abhishek Nagaraj, Michael Rose (discussant), Stephen Pinker, Rajiv Sethi, Carolyn Stein, Scott Stern, Toby Stuart, Valentina Tartari, Matteo Tranchero, and Josh Graff Zivin for their helpful comments, suggestions, and encouragement. We are grateful to the audiences at The Max Planck's RISE Workshop, University of Bordeaux, The Workshop in Economics of Innovation Complexity, Boston University's TPRI Seminar, DRUID, The University of California: Merced, The FIRE faculty conference, and The University of California, Berkeley Haas's Macro Research Lunch for their feedback. Portions of the coding and copy-editing were assisted by large language models. We thank Christian Venezuela for his outstanding research assistance. Corresponding author: please email kgulati@berkeley.edu.

"Science can flourish only in an atmosphere of free speech."
Albert Einstein (1930)

1 Introduction

Free speech has long been regarded as a cornerstone of democratic societies and a foundation for debate, innovation, and the pursuit of knowledge. Beyond its civic importance, open discourse also underpins organisational learning and innovation - processes that depend on the free exchange of ideas and differing views. Yet in modern professional environments, expressing one's views may entail professional risks. In 2020, for example, data scientist David Shor came under scrutiny after tweeting about an academic study in the *American Political Science Review* that analysed how violent versus nonviolent protests in the 1960s influenced voter behaviour. Despite the tweet pertaining to his professional expertise, it coincided with the Black Lives Matter movement, and so it was perceived by some as lacking sensitivity to the contemporary political context, prompting criticism and backlash (Yglesias, 2020). Shor's case illustrates a growing and broader phenomenon: *backlash arising from individuals expressing views that are considered controversial, unpopular, or diverge from the prevailing opinions within their professional communities*. This phenomenon - sometimes colloquially termed 'cancel culture' - has spanned across sectors, from academia to major tech companies such as Google, Apple, and Meta (Kessinger, 2024). For instance, a recent poll found that 54% of respondents were concerned about facing professional repercussions, including termination, if they freely expressed their opinions on social media (Harvard CAPS & Harris Poll, 2021). These examples illustrate the increased organisational tension between individual expression and workplace dynamics in contemporary professional environments.

However, despite the prominence of this phenomenon in both public and academic discourse, we lack systematic evidence on its professional consequences. This paper addresses this gap by posing the question: *How does backlash against an individual's speech - henceforth referred to as speech sanctioning events - affect their professional outcomes?* To the best of our knowledge, this is the first empirical paper to quantify the professional consequences of this phenomenon. In doing so, we use academia as an empirical setting in which professional interactions are unusually observable, measuring professional consequences through changes in peer recognition (citations to prior work) and individual research output (publications).

Ex ante, it is unclear whether speech sanctioning events should affect an individual's professional outcomes, and if so, in which direction, for two reasons. First, many organisations endorse norms that protect individuals' freedom to express their views in order to foster open exchange and encourage divergent thinking. Second, visibility-based mechanisms pull in the opposite direction. The common adage

that "all publicity is good publicity", suggests that, because speech sanctioning events are highly salient, heightened attention may elevate an individual's profile in ways that generate favourable downstream outcomes. Conversely, these events may also impose reputational costs among colleagues, within departments, and across professional networks, which may suppress peer recognition and individual productivity. As a result, whether these events ultimately have an effect on professional outcomes remains an open empirical question. We examine these competing forces in academia, our empirical setting, where both mechanisms are especially salient and observable. Academic communities often explicitly espouse norms of impartiality and universalism, most prominently articulated by Robert Merton, who argued that scientific claims should be evaluated independently of the personal attributes or views of their authors (Merton, 1973). Under these ideals, a scholar's public speech should not influence how their research is received. Yet professional evaluation frequently departs from this model, as social judgments and reputational concerns shape whose ideas gain recognition and how peers engage with them. Indeed, media attention - a requirement for inclusion in our sample - is associated with higher downstream citations (Mumtaz, 2025; Sofer, 2024; Phillips et al., 1991), potentially elevating a scholar's public profile in ways that spill over into scholarly impact.

To investigate this question, we build on the insight that professional outcomes are shaped not only by the intrinsic quality of individuals' work, but also by social processes that govern evaluation, coordination, and reputation within organisations. Specifically, in our context, academic research is frequently affected by forces that originate outside the research process itself. Guided by this perspective, we begin by examining how speech sanctioning events reshape peer-to-peer dynamics, influencing how others engage with and credit the affected scholar's work. Peer recognition is not a neutral process: characteristics such as status (Azoulay et al., 2014; Reschke et al., 2018; Bol et al., 2018; Teplitskiy et al., 2022; Leahey, 2004), gender (Bikard et al., 2025; Koffi and Marx, 2025; Koffi, 2025), and institutional or geographic origin (Fry and MacGarvie, 2024; Dumlao and Teplitskiy, 2025) have been shown to systematically influence how ideas are recognised and cited. Speech sanctioning events can operate through analogous channels, shaping how peers cite a scholar's intellectual contributions. We then turn to the production side of knowledge. Speech sanctioning events introduce a distinct kind of disruption to the process of knowledge creation. Similar external shocks, such as the loss of collaborators, have been shown to affect the production of new academic knowledge (see Azoulay et al., 2010; Oettl, 2012; Esposito, 2024; Mohnen, 2022; Samila et al., 2022). Likewise, speech sanctioning events can interrupt these processes, constraining scholars' productivity. By situating these events within the broader dynamics of knowledge production, recognition, and evaluation, we connect debates about free expression and speech sanctioning events to questions in the innovation and organisational literature, namely, how ideas are produced and recognised within professional communities.

Moreover, these events often trigger broader social reactions, which may prompt close peers in organisations either to distance themselves, or alternatively, to express support. These dynamics share conceptual parallels to those observed in scandals (see Luo and Zhang, 2022) - episodic, highly visi-

ble public events that arise from a discrete act provoking widespread moral outrage and reputational backlash (see Adut, 2018). In academia, scandals such as sexual misconduct or research retractions have been shown to reshape publication and citation patterns (Widmann et al., 2025; Collis, 2025; Lu et al., 2013; Azoulay et al., 2015). Yet, unlike the types of scandals examined in prior work, which typically involve clear violations of professional or ethical norms, the events we study do not necessarily entail misconduct or clear violations of ethical or professional norms, but rather stem from expressions that become publicly contested.

Academia as an empirical setting confers several advantages to study this phenomenon. It is a domain central to the creation of abstract knowledge that fuels innovation, technological progress, and long-run economic growth. This setting has long served as a natural laboratory for economists and management scholars, allowing them to examine not only how knowledge is produced and evaluated but also how social structures, institutional norms, and professional incentives shape that process (see Stern, 2004; Bikard et al., 2019; Teodoridis et al., 2019; Rahmandad and Vakili, 2019; Vakili et al., 2022; Hegde and Sampat, 2015; Stuart and Ding, 2006; Cavalli, 2022; Ayoubi et al., 2019). In addition to these conceptual foundations, academia provides a particularly fitting setting for our research question for several reasons. First, speech sanctioning events have been particularly prevalent in the discourse in U.S. academia. For instance, Frey and Stevens (2023) document a sharp rise in attempts to penalise scholars' speech, noting that 90 of the top 100 U.S. universities have faced such events. A recent survey of university faculty showed that 40% of liberal faculty, 56% of moderate faculty, and 72% of conservative faculty fear losing their jobs or reputations for something they say aloud or post online (Honeycutt et al., 2023), indicating that all scholars across the political spectrum feel vulnerable to this phenomenon.[1] Second, academia often explicitly espouses ideals of free speech and open dialogue (Resnik, 2008), making it an ideal environment for examining the tension between these professional ideals in theory and their consequences in practice. Indeed, Resnik (2008, p. 2) notes, "freedom of speech is one of science's most important norms." Finally, the organisational features of academic work - knowledge-intensive tasks, reputation-sensitive collaboration, project-portfolio management, and mentoring of students - parallel those of many other professional environments. As such, academia offers a setting whose dynamics generalise to broader managerial and professional contexts.

We draw on a novel dataset assembled by the Foundation for Individual Rights and Expression (FIRE), which documents over 1,000 U.S. cases since 2000 where academics faced calls for sanctions due to their speech. Because such events often unfold internally and rarely enter official records, this dataset enables a systematic study of episodes that would otherwise remain difficult to observe. We link these FIRE data on affected scholars to bibliometric records from OpenAlex (Priem et al., 2022), constructing a panel dataset comprising 351,248 paper-year and 5,776 scholar-year observations. Citations are particularly informative for two reasons. First, they capture elements of peer-to-peer dynamics that are

[1] It is important to acknowledge that, like many surveys, this one is subject to common limitations such as response bias, among others.

often unobservable in other organisational contexts. Second, because our empirical design focuses on papers published by an affected scholar before a speech sanctioning event, any subsequent decline in citations reflects deliberate shifts in peers' behaviour rather than differences in the underlying quality of the paper. Unlike most organisational contexts, academia provides, therefore, an unusually transparent environment for tracing how speech sanctioning events shape productivity and peer-to-peer evaluation, all while allowing us to hold the underlying quality of the work constant.

We employ a series of staggered difference-in-differences designs to evaluate the impact of these events, incorporating individual and time fixed effects to account for unobserved heterogeneity across scholars and temporal shocks common to all. Our empirical framework focuses on two dimensions of professional outcomes that capture both demand and supply channels of scholarly activity. On the demand side, we assess peer-to-peer professional recognition by tracking citations to a scholar's prior body of work after the event. On the supply side, we measure productivity through the number of new publications scholars produce after a speech sanctioning event. Importantly, recognition also feeds back into productivity: diminished citations can signal professional distancing, thereby constraining scholars' ability to generate new research. Guided by this framework, we document several key findings on how speech sanctioning events affect both scholarly productivity and recognition, and we further probe the mechanisms underlying these effects.

First, on the demand side, we track changes in peer behaviour by examining citations to the focal scholar's prior body of work after the event. We observe a 4.8% reduction in citations to the individual's prior body of work, relative to comparable papers published in the same journal, volume, and issue. This reduction indicates a measurable withdrawal of peer recognition. Unlike formal organisational sanctions such as termination or demotion, the responses we study operate informally through peer behaviour.

Second, on the supply side, we examine how involvement in speech sanctioning events affects individuals' productivity. We find that affected scholars publish 20% fewer academic papers after the event compared to a control group of scholars with similar characteristics. This decline in output is consistent with a broader withdrawal of professional support. Prior work shows that peers, teams, and professional networks play a central role in sustaining scientific productivity; when colleagues distance themselves or reduce engagement, subsequent performance often deteriorates (Mohnen, 2022; Azoulay et al., 2010; Cohen-Zada et al., 2024; Jaravel et al., 2018; Poege et al., 2025).

Third, we show that the impact of these events extends beyond the focal individual and affects their close collaborators. Co-authors who worked closely with the affected individuals also experience declines in their own research output following the event. On average, close collaborators publish roughly 12.5% fewer papers relative to matched controls. This spillover suggests that speech sanctioning events disrupt not only the productivity of the affected individual but also the functioning of the broader collaborators connected to them.

To shed light on the mechanism, we examine how peers adjust their behaviour in response to a speech sanctioning event. *Ex ante*, it is not clear if 'close' or 'distant' scholars in the focal scholars' academic network would be more likely to change their citation behaviour. We define close scholars as those who either work at the same institution or have coauthored with the focal scholar prior to the event; all others are classified as distant. Distant scholars might reduce citations simply because they lack strong collaborative or social ties. Conversely, close scholars may face greater professional or reputational concerns, paralleling patterns documented in the literature on scandals (Luo and Zhang, 2022), and thus may be more inclined to distance themselves. We test these competing explanations by comparing citation patterns between close and distant peers. The evidence supports the latter mechanism: scholars with close professional ties disproportionately reduce their citations to the focal scholar's prior work. This behaviour is consistent with professional distancing, suggesting that reputational concerns lead close peers to adjust their engagement to minimise their own exposure to controversy. This mechanism also helps explain the decline in the focal scholar's subsequent productivity because the affected scholars lose the professional support that normally facilitates continued scholarly work.

We further probe into the underlying mechanism by distinguishing between work-related and non-work-related speech sanctioning events, categorising cases based on whether the speech sanctioning events fall within or outside the scholar's domain of expertise. This distinction allows us to test competing explanations for the reduced citation patterns. If peers are rationally updating their beliefs about the focal scholar's work quality through Bayesian discounting, we would expect any citation effect to be concentrated within the group of work-related events. In such cases, speech sanctioning events directly tied to a scholar's research domain could reasonably prompt peers to reassess the credibility of their related academic contributions. However, if citation reductions extend beyond the scholar's field of expertise, this would provide stronger evidence that reputational concerns and social distancing are driving peer responses. Our findings support the latter interpretation: citation penalties arise for events both within and outside a scholar's domain of expertise, providing further evidence for the professional distancing mechanism.

Our study contributes to several literatures on organisations and innovation. First, we contribute to research on the organisational processes through which knowledge is evaluated and recognised. A large body of work shows that recognition is shaped not only by the intrinsic quality of ideas but also by social attributes such as status, gender, and institutional or geographic affiliation (Azoulay et al., 2014; Koffi, 2025; Fry and MacGarvie, 2024). We show that speech sanctioning events introduce a novel source of distortion into these evaluative processes: peers systematically adjust their recognition of a scholar's existing work following these events. By documenting how peer-to-peer recognition responds to these events, our findings highlight the role of informal social dynamics in governing evaluation within professional communities.

Second, we extend the innovation literature by identifying a previously unexplored organisational dis-

ruption to knowledge production. Prior work has focused on shocks such as collaborator loss, funding cuts, or institutional restructuring - events rooted in academic or organisational processes. In contrast, we study a disruption that originates in social and reputational conflict. We show that speech sanctioning events meaningfully reduce subsequent research output, illustrating how innovation is constrained not only by resource availability or team structure, but also by the social environment in which work is embedded. More broadly, our findings suggest that similar reputational disruptions may hinder individual productivity in other organisational and firm settings where collaboration, peer evaluation, and reputation are central to performance.

Third, our study broadens the scandals and organisational reputation literature by examining a form of controversy that stems from speech sanctioning incidents rather than formal rule-breaking. Whereas prior research focuses on what may be considered as 'illegitimate scandals', which involve fraud, harassment, or misconduct, we show that speech sanctioning events can generate comparable reputational fallout. Peers selectively withdraw recognition and engagement, revealing that reputational contagion can arise even when ethical or legal norms are not violated.

Taken together, while the phenomenon of speech sanctioning events dominates academic and public discourse, there has been little systematic evidence of its professional impact. As Clark et al. (2023, p. 1) note, "Science is among humanity's greatest achievements, yet scientific censorship is rarely studied empirically." To our knowledge, this study provides the first empirical analysis of how speech sanctioning events affect careers and peer behaviour. We show that these events not only diminish an individual's productivity and recognition but also reverberate through professional networks, as colleagues adjust their behaviour by professionally distancing themselves from the focal scholar.

Although we exploit academia as our setting, this study has broader implications. For managers and organisations in other settings, our findings highlight how speech sanctioning events can trigger tangible costs: reducing performance, weakening team cohesion, and altering professional ties and social dynamics, illustrating the need to understand and manage the organisational consequences of these events. Understanding how such reputational contagion unfolds has direct implications for organisations seeking to balance open expression, inclusivity, and performance amid an era of increasing ideological polarisation.

2 Related Literature & Conceptual Framework

Speech sanctioning events constitute a relatively new and understudied phenomenon. We study this phenomenon in academia, where professional interactions, recognition, and productivity are unusually observable. Although our empirical setting is academic research, the underlying outcomes we examine: peer-to-peer dynamics, individual productivity, and peer distancing are likely to operate in a wide

range of organisational contexts. Nevertheless, we draw on the literature on innovation to develop a conceptual framework that links speech sanctioning events to changes in peer recognition and individual productivity, providing theoretical guidance while allowing the data to speak to the net effects.

Two insights from prior research motivate this framework. First, academic recognition is inherently social. A scholar's ideas do not circulate on their merits alone; they are interpreted through the lens of status, identity, and institutional prestige. Second, research on innovation shows that the production of new knowledge is vulnerable to disruptions that originate outside the content of research itself - from collaborator loss to funding cuts - because knowledge creation depends on stable professional networks and the broader social environment.

We build on these insights by examining a novel class of disruption: speech sanctioning events, defined as cases in which scholars face public controversy over their speech rather than over their research. Unlike other shocks, these events do not alter the content or quality of a scholar's research. Instead, they perturb the social architecture within which recognition and research unfold. Conceptually, they sit at the intersection of three literatures: (1) how peers allocate attention and credit, (2) how external shocks disrupt knowledge production, and (3) how the effects of scandals propagate through networks.

Bringing these literatures together, we propose a framework in which speech sanctioning events reshape both the demand for a scholar's ideas (peer recognition, as reflected in citation behaviour) and the supply of new ideas (research productivity). This framework allows us to map how speech sanctioning events cascade through professional networks and to identify the channels through which such events generate professional costs.

2.1 Related Literature: Innovation

2.1.1 Peer Recognition as a Social Process

One strand of research within the innovation literature examines how ideas are recognised and attributed credit. This research emphasises that the impact of new ideas depends not only on their intrinsic quality but also on the social position and identity of their creators, the institutional environments in which they emerge and operate under, and the norms governing evaluation and reputation in academia. Status hierarchies and cumulative advantage processes play a central role in determining which ideas are cited or funded (Teplitskiy et al., 2022). For instance, Azoulay et al. (2014) and Bol et al. (2018) document "Matthew effects" in science, whereby recognition accrues disproportionately to already high-status scientists and institutions. Similarly, other work shows that gender shapes visibility and credit attribution (see, for example, Bikard et al., 2025; Koffi, 2025), while Fry and MacGarvie (2024) demonstrate that institutional origin systematically affects how academic contributions are received. Collectively, this literature reveals that recognition in science is, in part, social: ideas are not

judged in isolation, but through the lens of who produces them and where they originate.

Our study extends this literature by introducing a new source of variation in professional recognition: speech sanctioning events. These incidents do not alter the content or quality of a scholar's research, yet they can meaningfully shift how peers engage with their prior work. By linking these events to downstream citation behaviour, we show that peer recognition is sensitive to events that operate outside the research process. In doing so, we contribute to a growing literature in management and organisational theory on how identity, status, and social evaluation jointly shape the circulation of knowledge within expert communities.

2.1.2 Disruptions to the Production of New Knowledge

A large literature examines how new knowledge is produced, emphasising that new knowledge is not solely a function of individual talent or creativity, but is shaped by the social, institutional, and organisational environments in which academics work. Collaboration and team structure play a central role: complementarities in expertise, the division of labour, and cumulative team-specific capital are central to the production of new ideas (Bikard et al., 2019; Teodoridis et al., 2019; Jaravel et al., 2018; Wu et al., 2019). Other work has examined how institutional environments and incentive systems - such as funding mechanisms, tenure structures, and evaluation norms - shape which ideas are pursued and how they are developed (Stern, 2004; Hegde and Sampat, 2015; Vakili et al., 2022; Rahmandad and Vakili, 2019; Tripodi et al., 2025). Related research shows that organisational and network factors, from proximity to star scientists to exposure to diverse peers, further structure the generation of new ideas (Azoulay et al., 2010; Oettl, 2012; Truffa and Wong, 2025). Collectively, this work demonstrates that the academic production function is embedded within broader social and institutional architectures.

Building on these insights, a growing body of work examines how external shocks or disruptions - such as the loss of collaborators, funding changes, or institutional instability - affect knowledge production and career trajectories (Waldinger, 2012; Mohnen, 2022; Poege et al., 2025; Samila et al., 2022; Gulati and Helmers, 2025). These studies reveal that knowledge creation depends critically on the stability of professional networks. More broadly, this body of research illustrates that the production of new ideas is partly social: the institutions, norms, and peer structures surrounding scientists are not merely background conditions but active components of the research process.

Our study introduces a new type of disruption: speech sanctioning events. Unlike collaborator loss, funding shocks, or administrative restructuring - which stem from academic or institutional factors - these events originate in social and reputational conflict. We show that these socially rooted disruptions can meaningfully hinder the production of new knowledge, revealing a previously unexamined channel through which innovation can be constrained.

2.2 Related Literature: Scandals

Although the speech sanctioning events that we focus on in this paper may not immediately align with conventional notions of a scandal, they share many of the same underlying dynamics, which we draw on for our conceptual framework. A scandal is typically defined as an episodic, highly visible public event, often triggered by a discrete act that provokes widespread backlash or moral outrage (see Adut, 2018). The speech sanctioning events we examine exhibit similar properties: they are sudden, public, and reputationally consequential events that expose individuals to professional ramifications.

A growing literature on scandals demonstrates that the social consequences of scandals can be mediated by the structure and strength of social ties. For example, Luo and Zhang (2022) show that the Weinstein scandal and the broader #MeToo movement reshaped collaboration networks in Hollywood: producers previously connected to Weinstein became 35% more likely to hire female writers, with the strongest effects among female producers and those most closely linked to him. Similarly, Zheng and Wang (2025) find that the Facebook-Cambridge Analytica scandal generated heterogeneous responses among open-source developers depending on the intensity of their ties - some distanced themselves from Facebook projects, while others, particularly those with strong relational ties, intensified their contributions. Winkler et al. (2024) demonstrate that the R. Kelly sexual abuse scandal reduced music consumption mainly through supply-side forces, i.e., the removal of his songs from platforms, rather than shifts in consumer demand. However, after a popular documentary on R. Kelly was released covering the scandal, there was an *increase* in streams of his music, which came disproportionately from existing fans. Together, these studies reveal that the effects of scandals are mediated by social proximity: close ties can yield either avoidance or reinforcement, depending on whether association is perceived as reputationally costly or as an act of solidarity.

In academia, analogous patterns emerge. Scandals in the form of paper retractions or sexual harassment allegations have been shown to erode an individual's subsequent productivity and peer-to-peer recognition in the form of citations (Widmann et al., 2025; Collis, 2025; Lu et al., 2013; Azoulay et al., 2015).

Collectively, however, the existing literature has largely centred on what might be termed illegitimate scandals – those involving clear violations of legal or ethical standards. In such cases, the rationale for reputational damage and professional distancing is arguably expected or justified. Yet, this focus overlooks a growing category of incidents, such as speech sanctioning events, that emerge in the absence of any formal misconduct. This paper builds on the scandals literature by looking at this conceptual gap by examining a distinct class of scandals in which individuals are penalised not for breaking rules, but for expressing minority or controversial views that nonetheless do not breach legal and ethical norms.

2.3 Conceptual Framework

Because speech sanctioning events constitute a new and largely unexplored phenomenon, we deliberately refrain from formulating ex ante empirical predictions about their effects. Rather than specifying directional hypotheses, we draw on the related literatures above, on innovation and scandals, to inform our conceptual framework and empirical strategy. These literatures identify the mechanisms through which reputational shocks and peer responses can shape both individuals' recognition and production, providing theoretical guidance while allowing for a data-driven approach to reveal the net effects.

Building on these insights, we distinguish between two complementary channels through which speech sanctioning events may affect academic careers: the demand side, concerning the recognition of existing work, and the supply side, concerning the production of new knowledge. On the demand side, we examine how peers engage with the scholar's prior work, as reflected in citations accrued after the event to the scholar's prior body of work. On the supply side, we assess whether scholars continue to produce research after an event, measured through subsequent publications. Importantly, recognition and productivity are intertwined: a decline in citations can signal professional distancing, reducing access to opportunities and resources that facilitate the creation of new research.

2.3.1 Main Effects

Ex ante, the related literatures outlined above imply that the consequences of speech sanctioning events may be ambiguous: individuals involved in such events may experience positive, negative, or no effects on their professional outcomes, which, in the context of academia, translates to changes in publications produced and citations received. Below, we outline the competing mechanisms implied by prior research.

First, many professional communities, such as academia, are guided by norms and institutions that emphasise impartiality, objectivity, openness, and merit-based evaluation. The sociologist Robert Merton famously articulated a set of academic norms, including *universalism*, which holds that academic claims should be evaluated independently of the personal attributes or views of their authors (Merton, 1973). If these norms function as intended, they ought to buffer scholars from reputational shocks, implying that speech sanctioning events may have little to no effect on professional outcomes.

Consistent with these norms, one might expect that only clear ethical or legal violations would trigger professional penalties. Indeed, the existing scandals literature largely centers on such cases, often involving clear breaches of ethical or legal norms, such as fraud, sexual harassment, or gross professional misconduct, which can result in reputational harm and formal sanctions. By contrast, the speech sanctioning events we examine involve speech that remains within the boundaries of ethical and legal standards of professions and may even affirm central professional values such as open inquiry and

viewpoint diversity. Because these events concern expression rather than misconduct, there is no clear theoretical reason to expect systematic professional penalties.

Second, prior research has treated media-driven visibility as a consequential resource: media coverage shapes reputation and access to opportunities (Fombrun and Shanley, 1990; Deephouse, 2000; Rindova et al., 2007; Pollock and Rindova, 2003). The common adage, *"all publicity is good publicity"*, suggests individuals involved in speech sanctioning events could experience favourable career outcomes due to increased visibility. Inclusion in our analysis requires some level of media attention, be it local or national, which may raise awareness of the scholars and their work. In the academic context, prior research supports the idea that media exposure can generate positive career advantages (Mumtaz, 2025; Qiu et al., 2024; Phillips et al., 1991; Sofer, 2024). Applying this to our setting and research question, the literature on media coverage suggests that individuals may benefit from the heightened attention generated by involvement in a speech sanctioning event, possibly translating to greater output.

Even negative or contentious attention can generate positive outcomes. For instance, Berger et al. (2010) show that under certain conditions, negative product reviews can actually increase purchase likelihood by raising consumer awareness. Likewise, book banning in the US can lead to increased circulation (Ananthakrishnan et al., 2025). In the context of academia, Petersen et al. (2019) find that contrarian climate scientists - those who challenge the mainstream scientific consensus - receive disproportionately more media coverage than their consensus-aligned peers, despite comprising a small fraction of the field. In our setting, such visibility effects could translate into higher citations or increased professional engagement.

Furthermore, network dynamics may further amplify positive effects. Prior work shows that the individuals closely connected to someone embroiled in a controversy can mobilise on that person's behalf, helping to convert negative attention into professional gains (Zheng and Wang, 2025; Winkler et al., 2024). A related literature demonstrates that when prominent scientists die prematurely, their collaborators act as a "sales force" to sustain the visibility of their work (Azoulay et al., 2019). Analogously, speech sanctioning events may trigger solidarity among collaborators or allies, reinforcing the focal scholar's visibility and supporting their research trajectory.

On the other hand, speech sanctioning events may also generate negative professional consequences. The broader scandals literature shows that individuals publicly associated with controversy often face reputational penalties and diminished professional opportunities (Luo and Zhang, 2022). Parallel patterns appear in academic contexts: research on retractions and sexual harassment allegations documents sharp declines in citations, collaboration, and career advancement for implicated scholars (Widmann et al., 2025; Collis, 2025; Lu et al., 2013; Azoulay et al., 2015). These findings indicate that speech sanctioning events may hinder individuals' professional outcomes.

In sum, these mechanisms imply several distinct possibilities. First, outcomes may be negligible if meritocratic norms and evaluation prevail. Second, outcomes may be beneficial if controversy generates visibility that increases attention to the scholar's work. Similarly, outcomes may be beneficial if close collaborators or allies mobilise in solidarity, reinforcing support for the focal scholar. Third, outcomes may be adverse if the dynamics observed in the scandals literature apply. Because the theory from the prior literature does not clearly indicate which of these forces should dominate, we remain agnostic and let the data reveal the net effect.

2.3.2 Mechanism

To investigate the mechanism, we look at *which type* of peers adjust their behaviour after the event. We classify peers by tie strength - either through prior collaboration or contemporaneous institutional affiliation - and compare how their behaviour towards the focal individual's pre-event work changes after the speech sanctioning events. This approach isolates which relationships drive the response and maps the social and professional dynamics the event sets in motion.

Ex ante, it is not clear whether *close* or *distant* peers are most likely to change their behaviour following a speech sanctioning event. On the one hand, weak-tie colleagues - those more *distant* from the focal individual - may disengage more readily because they have less relationship-specific investment. On the other hand, *close* collaborators may distance themselves to mitigate perceived reputational spillovers.

To elaborate, there are several reasons to expect close collaborators to come to the aid and support their colleague because of a speech sanctioning events. First, those with the strongest ties often have the most at stake: shared projects, co-authored work, and intertwined reputations give them incentives to support the focal scholar. Second, as shown in the scandals literature, close associates sometimes respond to controversy by reinforcing rather than severing connections. Zheng and Wang (2025) and Winkler et al. (2024), for example, find that individuals closely linked to a scandal can consolidate their relationships in its aftermath. Finally, Azoulay et al. (2019), for example, find that following the premature death of a superstar scientist, their close collaborators mobilised as a 'sales force' to promote the scientist's work. Analogously, when facing calls for speech-based sanctions, the close network of affected individuals might similarly come to the aid of their peer by providing additional support.

Conversely, concerns over reputational risk may dominate. Those most closely associated with the focal individual might seek to protect their own professional standing by disengaging or reducing visible association. After the Weinstein scandal, for example, Luo and Zhang (2022) find that it was the producers closest to Weinstein who were more likely to change their behavioural practices. This is consistent with prior work on scandals such as sexual harassment in academia, where close peers are disproportionately more likely to cite the focal scholar less (Widmann et al., 2025).

Taken together, these mechanisms suggest whether close or distant peers will adjust their behaviour more. This motivates our empirical test of how different types of ties respond.

3 Data & Descriptive Statistics

3.1 Data

We leverage a novel dataset from the Foundation for Individual Rights and Expression (FIRE), a non-partisan, non-profit organisation whose stated mission is to "defend and sustain the individual rights of all Americans to free speech and free thought."[2] In recent years, FIRE has led a systematic effort to document incidents in which scholars expressed their views and subsequently faced calls for professional sanction. The resulting dataset, *Scholars Under Fire* (Frey and Stevens, 2023), to which we obtained early access through a collaboration with FIRE, provides the most comprehensive record to date of speech sanctioning events in U.S. higher education, covering both public and private institutions from 2000 to the present. FIRE defines these events as "campus controversies with efforts to investigate, penalise, or otherwise professionally sanction a scholar for engaging in constitutionally protected forms of expression."(Frey and Stevens, 2023).

FIRE operationalises a "scholar" as "any individual who engages in acts of scholarship within the academic domain and has an official affiliation with a college or university. This includes teaching at a college or university, conducting research and submitting the findings to the peer-review process, and/or discussing peer-reviewed scholarship at professional academic events (for example, conferences, panel discussions)."(Frey and Stevens, 2023). Importantly for our analysis, FIRE excludes "deans and other administrators who have never held a faculty position, as well as researchers working for non-university-affiliated organisations," (Frey and Stevens, 2023), ensuring that our dataset focuses exclusively on individuals with academic appointments.

As of January 2023, FIRE's dataset contains 844 scholars.[3] We review each case to ensure it meets the scope of this study. First, we include only speech sanctioning events in which individuals directly expressed their own views.[4] Second, we remove events involving illegal conduct, such as potential sexual harassment. Third, we require that the event received media coverage - at the local or national level - since, without reaching the wider academic community, such events are unlikely to generate measurable professional effects. After these steps, 349 scholars are removed from the sample.

[2] https://www.thefire.org/about-us.

[3] Some scholars are involved in multiple events, and so we only take their first entrance into the dataset.

[4] For example, a Professor of Music at the University of Michigan was forced out of class for showing a 1965 movie of Othello in which actor Laurence Olivier darkened his skin to play Othello. Students later called for sanctions. We deem examples like this outside the scope of the paper as they are not a clear expression of the scholar's opinion.

For the remaining 495 individuals, we collect comprehensive bibliometric data using OpenAlex (Priem et al., 2022), an open-source bibliometric database that provides publication and citation information. To ensure data accuracy, we manually search the OpenAlex API for each scholar using a systematic four-step process: (1) searching by name and institutional affiliation, (2) identifying one of their publications and saving the publication ID, (3) using this publication to locate the scholar's unique OpenAlex ID, and (4) extracting their publication and citation information.

To enable meaningful pre- and post-event comparisons, we restrict our sample to scholars with at least one publication both before and after the speech sanctioning event. This requirement ensures that we can observe changes in productivity over time. Excluding scholars who may have left academia following the events in which they were involved prevents us from conflating the treatment effect with potentially voluntary exit decisions. This means that our estimates are conservatively estimated insofar as we are capturing only individuals who continue to publish. These steps decrease our sample to 301 scholars. This is our preliminary dataset.

Given the breadth of what constitutes a speech sanctioning event, the types of events in our dataset are fairly heterogeneous. To the best of our ability, we attempt to categorise the events to leverage the richness of the data. A significant portion of the data involves incidents related to socio-political issues both within and outside the scholars' fields of expertise. These incidents include statements, such as, critiquing or defending the Trump presidency, expressing opinions on the Israel/Palestine conflict, supporting or criticising Black Lives Matter or Blue Lives Matter movements, questioning Covid-19 policy responses, challenging diversity, equity, and inclusion initiatives in universities, feminist perspectives on transgender rights, critiques of gender theory, discussions of racial issues, and interpretations of religious matters, among others. The remainder of the data contains events within academic fields, including debates over legal justifications for interrogation techniques, criticisms of social justice activism in academia, discussions of eugenics, research on the male variability hypothesis and its implications for scientific performance, climate change discourse, and work on predictive policing, among others.

The incidents in our dataset span a wide range of communication contexts. Scholars expressed the relevant views in classroom discussions, published articles, social media posts, conference presentations, emails to colleagues, podcasts, and public lectures, among other settings.

To ensure our findings are not driven by particularly 'extreme' incidents, we exclude cases involving statements classified as hate speech, slurs, or other forms of offensive expression (see Appendix Section D for details on this classification). In such circumstances, one could argue that social distancing may be normatively justified, making these episodes less informative for identifying peer-driven distancing dynamics in the broader population of cases. For this reason, our preferred specification omits these observations. We also exclude cases in which scholars were ultimately fired, terminated, or resigned in connection with the incidents. These episodes involve institutional, legal, and administrative

processes that differ systematically from peer-driven social responses, thereby risking a conflation of organisational discipline with the informal distancing we seek to study. Importantly, our results remain robust when both types of these cases are included (see Appendix Section J)

3.2 Descriptive Statistics

Our descriptive statistics are organised into two sections corresponding to our main analyses and dependent variables: Section 3.2.1, the citation analysis, where the data is analysed at the paper level, and Section 3.2.2, the publication analysis, where the data are aggregated at the scholar level.

3.2.1 Descriptive Statistics: Citations

Appendix Section A Figure A1 plots the number of scholars affected by speech sanctioning events each year and compares those in the preliminary dataset to those used in the citation analysis. The final sample includes 173 treated scholars.[5]

Table 1 Panel A, presents descriptive statistics for the number of yearly citations, comparing treated and control groups before and after treatment. As described in more detail in Section 4, the treatment group consists of papers published by the affected scholars before the event; for each treated paper, we sample up to ten control papers published in the same journal, volume, and issue. This matching strategy controls for journal-, volume-, and issue-specific citation dynamics and relies on the assumption that, absent an event, papers published in the same outlet, volume, and issue would follow parallel citation trajectories and are comparable in underlying quality (see, for example, Azoulay et al., 2015).

The resulting dataset used in the citation analysis contains 5,182 treated papers and 35,076 control papers, for a total of 351,248 paper-year observations. Appendix Section E details the data cleaning and construction steps that produce this sample. As shown in Table 1 Panel A, both groups experience fewer citations in the post-treatment period relative to the pre-treatment period, which is expected given that, on average, papers tend to become less relevant as new or updated insights are generated. However, the decline is more pronounced for the treated group than for the control group: the treated group experiences a 9.2% drop, while the control group 5.2%, suggesting treatment-specific effects beyond natural citation ageing patterns.

Table 1 Panel B, presents descriptive statistics for the pre-treatment characteristics of affected scholars. The data reveals that most events occurred in recent years, with a mean treatment year of 2017.

[5] As a robustness check, we also estimate all results using an expanded sample that includes scholars involved in extreme events and those who were ultimately fired, terminated, or resigned due to the speech sanctioning events. In that specification, the number of treated scholars increases to 239. See Appendix Section J.

We calculate publication and citation stock as the discounted sum of an academic's output up to the treatment year, applying a standard 15% annual discount rate. Age at treatment is defined as the difference between the treatment year and the scholar's first publication year. The distribution of key metrics: citations per year, citation stock, publications per year, and publication stock exhibits the characteristic right skew typical of bibliometric data. To mitigate the potential influence of outliers, we restrict our analysis to the 99th percentile of the dependent variable.

To conclude, Figure 1 shows the evolution of average annual citations for treated and control papers. In the pre-event periods between -5 and -1 both groups exhibit relatively stable citation trajectories. Following the event, both groups experience a decline in citation rates, though the treated group experiences a more pronounced decline. These descriptive patterns are consistent with the estimates reported in Table 1 Panel A, providing model-free evidence of a decline in citations.

Table 1: Summary and Descriptive Statistics

Panel A: Summary Statistics of the Yearly Citations per Paper

Group	Period	Obs	Mean	SD	Median	Min	70th	80th	90th	95th	Max
Control	Pre	194,706	2.87	4.94	1	0	3	4	8	12	39
Control	Post	111,261	2.72	4.93	1	0	2	4	7	12	39
Treatment	Pre	28,808	3.16	5.23	1	0	3	5	9	13	39
Treatment	Post	16,473	2.87	5.03	1	0	3	4	8	13	39

Panel B: Yearly Descriptive Statistics of Treated Scholars (Pre-Treatment)

Variable	Mean	SD	Med.	Min	70th	80th	90th	95th	Max
Treatment year	2017	3.8	2019	2001	2020	2021	2021	2021	2021
Citations per year	71	204	6	0	32	72	167	339	2,547
Citation stock	311	948	23	0	127	281	727	1,417	13,102
Publication per year	1.85	3.12	1	0	2	3	5	7	49
Year of first publication	1991	13	1990	1965	2000	2004	2009	2012	2019
Age at treatment	26	13	26	1	34	38	44	47	51
Cumulative publications	29	52	13	1	28	42	64	98	563
Publication stock	10	15	5	0	10	13	21	31	187
Citations over publication stock	21	66	6	0	15	26	48	79	1,575

Notes: **Panel A** reports paper-level citation summary statistics, categorised by treatment status and period. The unit of observation is a paper-year. To limit the influence of extreme values, we exclude yearly citation counts above the 99th percentile (39 citations). **Panel B** reports scholar-level characteristics for treated individuals prior to their speech sanctioning event. Stock variables (citation stock and publication stock) are constructed as discounted sums of all prior output up to the treatment year, applying a 15% annual discount rate. Age at treatment is defined as the difference between the treatment year and the scholar's first publication year. All variables in Panel B are aggregated at the individual scholar level. For each variable, we report the mean, standard deviation, median, minimum, the 70th, 80th, 90th, and 95th percentiles, and the maximum.

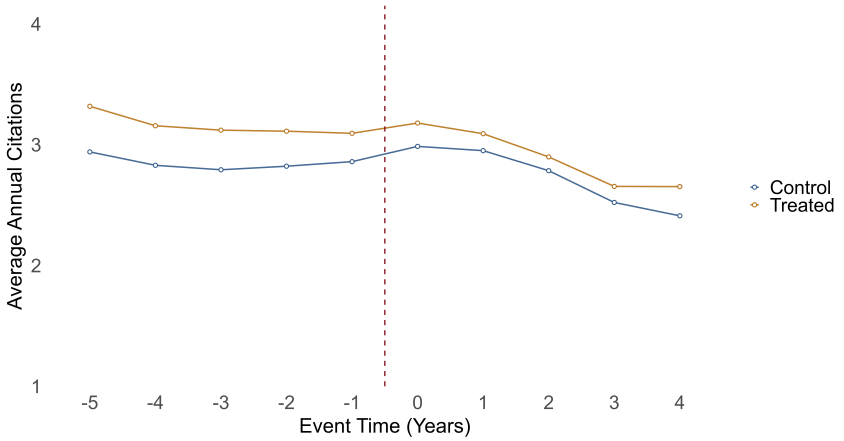


Figure 1: This figure presents the average annual citations to pre-event publications for treated papers (yellow) and matched control papers (blue), plotted relative to the year of the speech sanctioning event (t = 0).

3.2.2 Descriptive Statistics: Publications

Appendix Figure A2 plots the number of treated scholars in our publication sample over time. Due to OpenAlex API limitations that prevent us from retrieving control-group publications beyond a 10-year window, we restrict the publication analysis to scholars whose events occurred in 2012 or later.[6] Although this reduces historical coverage, the resulting sample more accurately reflects contemporary academic and social dynamics. The final publication-level dataset consists of 77 treated scholars and 466 matched control scholars, totalling 5,776 scholar-year observations. Appendix Section F details the data cleaning and sample-construction steps underlying this dataset. Consistent with the citation analysis, we exclude cases classified as extreme as well as cases in which the scholar was ultimately fired, terminated, or resigned in connection with the incident. When including these cases, the sample includes 104 treated and 649 control scholars, respectively, and the results are robust (see Appendix Section J). Control scholars are matched on the following criteria, measured in the year prior to the event: (1) the primary topics of a scholar's work, (2) geographic location, (3) number of publications, (4) number of citations, and (5) academic age.

Table 2 Panel A presents descriptive statistics for the number of yearly publication counts across treated and control groups in the pre- and post-treatment periods. The treated group exhibits a 17.5% decline in average publications following the event, while the control group maintains relatively stable publication rates over the same timeframe.

Figure 2 illustrates the raw publication trajectories of treated and control scholars before and after the event year. Prior to the event, the two groups track each other closely, consistent with parallel pre-

[6] By contrast, the citation analysis includes the full set of affected scholars dating back to 2000, as it does not rely on the same matching constraints.

trend patterns. Following the event, a visible divergence emerges, with treated scholars exhibiting a sharper decline in yearly publication counts. Because these values are unadjusted, the figure should be interpreted solely as descriptive evidence. It is important to note that the potential convergence in year 4 should be treated with caution because this has fewer observations, reflecting the concentration of incidents in more recent years. Taken together, Table 2 Panel A and Figure 2 provide initial, model-free evidence consistent with a post-event decline in publication output among treated scholars.

Table 2 Panel B reports key demographic and academic characteristics, as of the treatment year, for scholars in the treated group and their matched controls, including treatment year, publication counts, citation metrics, and co-authorship patterns, showing that they have comparable observable characteristics.

Table 2: Summary and Descriptive Statistics

Panel A: Summary Statistics of the Number of Yearly Publications

Group	Period	Obs	Mean	SD	Med.	Min	70th	80th	90th	95th	Max
Control	Pre	3,285	4.50	5.00	3	0	5	8	12	15	32
Control	Post	1,695	4.45	5.38	2	0	5	7	12	16	32
Treatment	Pre	530	4.31	5.02	3	0	5	7	11	15	31
Treatment	Post	266	3.56	4.07	2	0	4	6	8	13	31

Panel B: Descriptive Statistics of Treated and Control Scholars at Treatment

Group	Variable	Mean	SD	Med.	Min	70th	80th	90th	95th	Max
Treated	Treatment year	2018	2	2019	2013	2020	2020	2021	2021	2021
	Publications	84	82	55	1	102	142	217	245	398
	Citations	1,543	2,476	447	0	1,518	2,731	4,391	6,496	14,207
	First publication	1988	15	1991	1950	1998	2001	2005	2011	2015
	Coauthors	105	166	42	0	101	180	288	364	1,100
Control	Treatment year	2018	2	2019	2013	2020	2020	2021	2021	2021
	Publications	89	85	64	1	107	142	213	243	662
	Citations	1,492	2,304	571	0	1,623	2,588	4,158	5,885	28,441
	First publication	1990	14	1992	1950	1999	2003	2008	2011	2019
	Coauthors	144	222	67	0	157	224	352	556	2,535

Notes: **Panel A** presents yearly publication statistics for treated and control scholars before and after the event. **Panel B** reports scholar-level characteristics measured at the treatment year, including mean, median, and selected percentiles (70th, 80th, 90th, 95th). All statistics are computed at the individual scholar level, making the unit of observation a scholar-year in Panel A and a scholar in Panel B. Treatment year refers to the year the scholar was involved in the speech sanctioning event. Publications refers to the number of publications a scholar has at treatment. Citations refers to the number of citations a scholar has at treatment. Coauthors refers to the number of coauthors a scholar has at treatment.

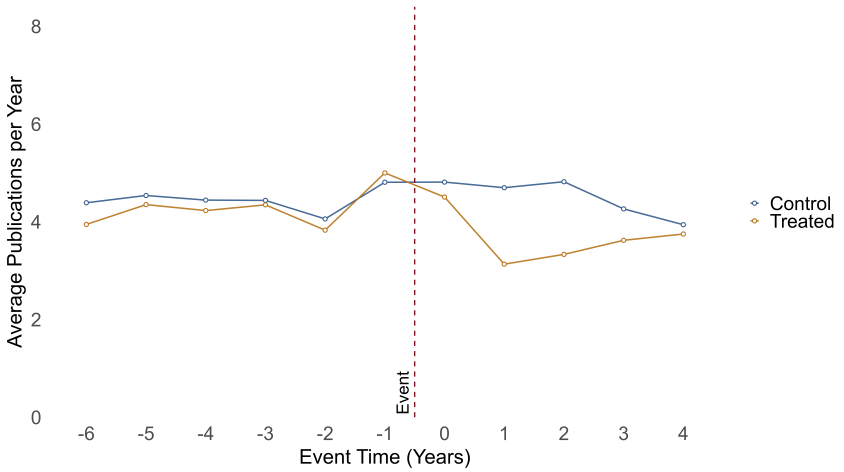


Figure 2: This figure presents raw averages of yearly publication counts. $t = 0$ marks the year of the speech sanctioning event (red dashed line). Treated scholars (yellow) are compared to matched control scholars (blue). Because the values are unadjusted, the plot serves as descriptive evidence only and does not incorporate fixed effects or covariate adjustments used in the main specification.

4 Method & Results

We first analyse the peer-recognition effects, focusing on how citations to a scholar's prior work change after the event in Section 4.1. We then examine the productivity effects, assessing how the event influences the scholar's subsequent publication output in Section 4.2.

4.1 Effect on Citations

To test whether the peer recognition changes following the speech sanctioning events, we employ a difference-in-differences design that examines citations to scholars' pre-event publications. By focusing on work published before the speech sanctioning events, we can isolate changes in peer behaviour while holding research quality constant. The control group consists of up to 10 randomly selected papers from the pool of papers published in the same journal, volume, and issue as the treated paper. This approach controls for the citation trajectories of the papers and assumes that, absent an event, papers published in the same journal, volume, and issue follow, on average, similar citation trajectories and are of comparable quality (see Azoulay et al., 2015), for example, with a similar journal-volume-issue approach). We estimate the following specification:

$$\text{Citations}_{it} = f\left(\beta_1 \text{Treat}_i * \text{Post}_{it} + \Gamma_i + \varphi_t + \eta_k\right) + \varepsilon_{it} \tag{1}$$

where the dependent variable Citations_{it} represents the number of citations received by paper i in year t. Given the count nature of our dependent variable, we estimate the model using a Poisson regression.[7] The binary treatment indicator Treat_i equals one for papers authored by treated scholars and zero otherwise. Note that the affiliation with the treatment or control groups (Treat_i) is time-invariant and, hence, included in the paper's fixed effect (Γ_i). Γ_i controls for inherent differences between papers caused by unobservable factors in the form of paper fixed effects. The binary variable Post_{it} switches to one beginning the year after the treated scholar is involved in a speech sanctioning events. Year fixed effects (φ_t) capture common temporal trends affecting all publications, while η_k represents journal-publication year fixed effects and controls for time-varying, journal-specific factors by interacting the publication year with journal identifiers. The standard errors are clustered at the article level. The main result of the model is provided by the coefficient β_1, which captures the average difference in the change of citations between treatment and control papers after the event. To account for the staggered effect of the treatment, we employ the staggered DiD methodology developed in Callaway and Sant'Anna (2021).[8].

Figure 3 shows an event study demonstrating the parallel movement of the dependent variable in the pre-treatment periods between treatment and control groups. The specification in Figure 3 extends Equation (1) by interacting the year dummies (φ_t) with the treatment indicator (Treat_i). To account for the staggered nature of the speech sanctioning events in our dataset, we estimate the model using the method proposed by Callaway and Sant'Anna (2021). As shown in Figure 3, there is no significant difference between the two groups prior to the events.

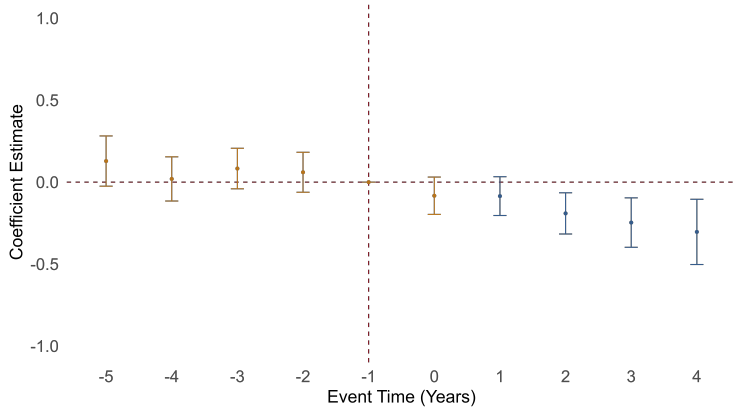


Figure 3: Event-study plot using the Callaway and Sant'Anna (2021) difference-in-differences estimator, which accounts for staggered treatment timing and cohort heterogeneity. Points represent period-specific treatment effects relative to the year before the event ($t = -1$), and the vertical lines show 95% confidence intervals. Estimates for $t < 0$ assess pre-trend behaviour, whereas coefficients for $t \geq 0$ capture the dynamic impact of speech sanctioning events on subsequent citation outcomes.

[7] The results are robust to alternative specifications, including negative binomial models, and OLS regressions. See Appendix Section I.

[8] The findings are robust to alternative difference-in-differences estimators, including Sun and Abraham (2021) and conventional two-way fixed effects, see See Appendix Section I.

Table 3 reports the results of Equation (1). Columns (1)-(4) present a sequence of specifications gradually including additional fixed effects. Column (1) reports the baseline model without fixed effects. The sample here is significantly larger because adding article fixed effects forces the exclusion of articles with no within-article variation in the dependent variable. Column (2) adds article fixed effects, thereby absorbing all time-invariant characteristics of each paper. Column (3) further incorporates year fixed effects, which control for annual fluctuations in overall citation practices. Finally, Column (4) includes all fixed effects as indicated in Equation (1). Across these specifications, the estimated post-event treatment effect remains negative and highly stable, underscoring the robustness of our findings. As mentioned, given that Treat_i is time-invariant, it is collinear with the article fixed effect and therefore it is excluded from the regression in columns (2), (3) and (4). The coefficient for the interaction term $\text{Treat}_i * \text{Post}_{it}$ is negative and statistically significant ($p < 0.001$), providing evidence of a decrease in citation trajectories for the treatment group. More specifically, treated papers receive approximately 4.8% fewer citations compared to the control group after their authors are involved in a speech sanctioning events.

Table 3: The Impact of Speech Sanctioning Events on Citations

	(1)	(2)	(3)	(4)
Dependent Variable		Citations		
Treat	0.0955***			
	(0.0221)			
Post	-0.0432***	-0.0242***	0.0235***	0.0235***
	(0.0057)	(0.0045)	(0.0066)	(0.0067)
Treat × Post	-0.0452***	-0.0464***	-0.0491***	-0.0491***
	(0.0156)	(0.0128)	(0.0127)	(0.0128)
Fixed Effects:				
Article	No	Yes	Yes	Yes
Year	No	No	Yes	Yes
Journal × Publication Year	No	No	No	Yes
Model:		Poisson		
S.E.: Clustered		Article		
Observations	682,943	351,266	351,248	351,248
Squared Correlation	0.00143	0.77914	0.78122	0.78122
Pseudo R^2	0.00216	0.57595	0.57695	0.57695
BIC	4,121,973.1	1,623,735.2	1,621,401.9	1,680,280.9

Notes: This table reports difference-in-differences estimates of how a speech sanctioning event affects a scholar's annual citation count. The coefficient of interest, *Treat* × *Post*, captures the change in citations after the event relative to the control group. Columns (1)-(4) present specifications with progressively additional fixed effects. Significance levels: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\dagger}p < 0.1$.

Taken together, these results show that speech sanctioning events meaningfully disrupt the citation trajectories of a scholar's prior body of work, in part because peers adjust their behaviour towards the

affected individual. Such dynamics stand in tension with the Mertonian norm of universalism, which suggests that academic contributions should be evaluated independently of the personal attributes or circumstances of their authors (Merton, 1973). Instead, our findings align with a growing literature demonstrating that the recognition of academic ideas depends not only on their intrinsic quality but also on social factors such as status, gender, and institutional or geographic origin (Azoulay et al., 2014; Koffi, 2025; Fry and MacGarvie, 2024).

4.2 Effect on Individual Productivity

Our second main analysis examines the impact of speech sanctioning events on the affected individuals' scholarly productivity. Our treatment group comprises scholars who experienced speech sanctioning events between 2012 and 2021. The control group consists of scholars with comparable observable characteristics, measured in the year prior to each event. We employ a matching procedure based on five criteria: (1) primary research field[9], (2) location, (3) publication count, (4) citation count, and (5) academic age. To ensure close matches, we match each treated scholar with controls whose publication counts fall $\pm$ 20% and citation counts $\pm$ 30% of the treated scholar. Academic age, defined as the number of years between the first publication and the treatment, instead, is set to a maximum of 6 years difference between treated and control scholars, i.e., the 1st quartile of the distribution across all plausible control scholars.

Using this matched sample, we estimate the following difference-in-differences model:

$$\text{Publications}_{st} = f\left(\beta_1 \text{Treat}_s * \text{Post}_{st} + \Gamma_s + \varphi_t\right) + \varepsilon_{st} \tag{2}$$

where the dependent variable Publications_{st} is the number of publications produced by scholar s in year t. Treat_s is a binary variable that equals one for scholars in the treatment group and zero for control group scholars. Since this assignment is time-invariant, it is absorbed by the scholar's fixed effect (Γ_s), which controls for unobservable characteristics that may influence publication productivity. The variable Post_{st} is a binary indicator that switches to one in all years following a treated scholar's speech sanctioning event. Year fixed effects (φ_t) capture common temporal trends affecting all scholars. The standard errors are clustered at the author and treated scholar levels. β_1 is the coefficient of interest, and we estimate the model using a Poisson regression to account for the count nature of our dependent variable.

Figure 4 presents the event study using Callaway and Sant'Anna (2021). It shows stable parallel trends

[9] This is the top-level 1 and 2 concepts from OpenAlex.

in the pre-treatment period.[10]. Following the event, a negative treatment effects emerge one year after the event.

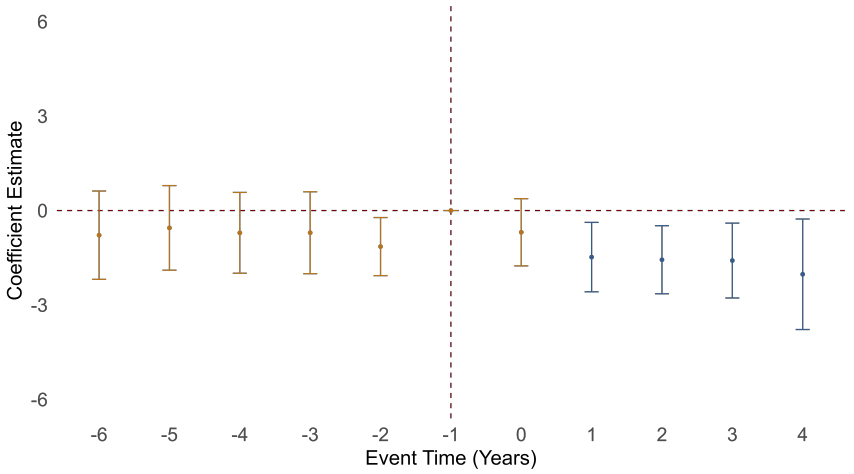


Figure 4: Event-study plot using the Callaway and Sant'Anna (2021) difference-in-differences estimator, allowing for heterogeneous treatment timing and dynamic treatment effects. Each point reports the period-specific average treatment effect relative to the year before the event ($t = -1$), with 95% confidence intervals. Estimates for $t < 0$ assess pre-trend behaviour, whereas coefficients for $t \geq 0$ capture the dynamic impact of speech sanctioning events on subsequent publication outcomes.

Table 4 reports our main results based on estimating Equation (2). Column (1) presents the baseline model without any fixed effects. Column (2) introduces author fixed effects, absorbing time-invariant differences across scholars. Column (3) further adds year fixed effects. Across all model specifications, the coefficient on the interaction term $\text{Treat}_i * \text{Post}_{it}$ indicates a robust negative treatment effect: scholars experiencing speech sanctioning events publish approximately 20% fewer papers compared to their matched counterparts.

[10]The results are robust to alternative estimators as shown in Appendix Section I

Table 4: The Impact of Speech Sanctioning Events on Publications

	(1)	(2)	(3)
Dependent Var:	Publications		
Treat	-0.0392		
	(0.0896)		
Post	-0.0071	0.0033	0.0709
	(0.0343)	(0.0340)	(0.0584)
Treat × Post	-0.1843*	-0.2124*	-0.2201*
	(0.0910)	(0.0955)	(0.0967)
Fixed Effects:			
Author	No	Yes	Yes
Year	No	No	Yes
Model	Poisson		
S.E.: Clustered	Author & Treated Scholar		
Observations	5,776	5,776	5,776
Squared Correlation	0.00147	0.54654	0.55413
Pseudo R^2	0.00119	0.36713	0.37184
BIC	44,113.9	32,650.1	32,589.7

Notes: This table reports estimates of how scholars' annual publication output changes following a speech sanctioning event. The coefficient of interest, *Treat* × *Post*, captures the post-event change for treated scholars relative to their matched controls. Column (1) has no fixed effects, Column (2) includes author fixed effects, and Column (3) includes author and year fixed effects. Standard errors are clustered at the author and treated-scholar level. Significance levels: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\dagger}p < 0.1$.

It is important to note that the results of our two main analyses complement each other. The citation penalty observed in Section 4.1 helps to explain why affected scholars subsequently publish less. Citations are not only a measure of recognition but also a form of professional currency that shapes reputation, collaboration opportunities, and access to resources. When peers deliberately reduce citations to a scholar's prior paper, it weakens the scholar's standing within their academic network. As prior research shows, peers and teams are critical drivers of scientific productivity, both by conferring legitimacy and by facilitating collaboration. In this way, the observed citation effect helps to account for the decline in future publications.

5 Mechanism: Professional Distancing

To investigate the underlying mechanism driving these citation patterns, we examine which types of peers are more responsive in their behaviour to the individuals involved in the speech sanctioning events. Drawing upon the scandals literature discussed above, we test whether citation responses differ systematically between scholars with *close* versus *distant* professional relationships to the affected

individual.

We operationalise scholarly proximity using two complementary measures that capture different dimensions of professional closeness, i.e., more likely to share social and professional ties. First, we identify collaborative proximity through co-authorship networks, classifying scholars who previously co-authored with the affected researcher as 'close'. Second, we identify institutional proximity by examining citations from scholars affiliated with the same institution as the affected researcher at the time of the event. Finally, to reassure our confidence in the mechanism, we restrict our close coauthors' analysis to citations only originating from within the focal scholars' research field, ensuring that all the citing scholars are likely aware of the event, given the media coverage and field-level attention these events typically receive.

To examine this effect, we estimate the following equation:

$$\text{Citations}_{it} = f\left(\beta_1 \text{Close}_c * \text{Post}_{it} + \delta\Gamma_i + \varphi_t + \eta_k\right) + \varepsilon_{it} \tag{3}$$

where Citations_{it} indicates the number of citations received by paper i in year t. Close_c is a dummy variable that equals one when citations originate from close scholars and zero otherwise. The standard errors are clustered at the paper level. All other variables follow the same definitions as in Equation (2) above.

Table 5 reports the estimation results for both proximity measures. Column (1) reports the institutional results and highlights that scholars affiliated with the same institution as the focal scholar as of the year of the speech sanctioning events, tend to cite the focal scholar's prior work 25% less than scholars affiliated elsewhere. Similarly, in Column (2), we show that scholars who coauthored with the focal scholar before the event tend to cite the focal scholar's prior work 53% less than scholars who did not coauthor before the event. In Column (3), we take the sample used in Column (2) and further restrict the yearly citation count to papers coauthored by at least one scholar operating in the same field as the treated scholar and find a 63% drop.

Table 5: Social Distancing Mechanism: Institutional and Collaborative Proximity

Dependent Variable:	Institution citations	Coauthor citations	Coauthor citations (same field)
Post	0.0321	0.0407*	0.0127
	(0.0197)	(0.0190)	(0.0219)
Close	-3.242***	-4.733***	-4.619***
	(0.0378)	(0.0843)	(0.0938)
Post × Close	-0.2832***	-0.7548***	-0.9927***
	(0.0449)	(0.1378)	(0.1223)
Fixed Effects:			
Article	Yes	Yes	Yes
Year	Yes	Yes	Yes
Journal ID × Publication Year	Yes	Yes	Yes
Model		Poisson	
S.E.: Clustered		Article	
Observations	76,662	90,444	71,988
Squared Correlation	0.79849	0.80143	0.80329
Pseudo R^2	0.66859	0.70366	0.69059
BIC	247,266.9	260,652.5	203,108.8

Notes: This table reports difference-in-differences estimates of how speech sanctioning events affect citations from scholars who are *close* versus *distant* to the focal individual. *Close* is defined by: (i) institutional proximity (Column 1), measured as scholars affiliated with the same institution; (ii) collaborative proximity (Column 2), measured as prior coauthorship; and (iii) collaborative proximity within the same research field (Column 3). All specifications include article, year, and Journal ID × Publication Year fixed effects, and standard errors are clustered at the paper level. Significance levels: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\dagger}p < 0.1$.

Overall, these findings provide evidence for professional distancing as a mechanism. The fact that both institutionally proximate scholars and prior co-authors exhibit disproportionately larger citation reductions suggests deliberate behavioural changes. Close scholars likely face greater reputational risks from continued association and may face pressure to distance themselves from the focal scholar who is embroiled in a scandal. By reducing citations to controversial colleagues, these scholars can signal professional distance while minimising potential career costs. This strategic behaviour aligns with the literature on scandals.

Beyond our empirical results, survey evidence from FIRE corroborates the presence of professional distancing as a mechanism. In a recent survey of 209 scholars involved in speech sanctioning incidents, 47% reported losing professional relationships, and 40% reported experiencing workplace shunning (Honeycutt and Stevens, 2025). This supports the mechanism that we document empirically.

The evidence of professional distancing provides an additional channel for understanding why scholars face reduced publication output. Our analyses show that it is the colleagues most closely connected

to the affected scholars - former co-authors and institutional peers - who disproportionately reduce their citations to the affected scholar. When close peers withdraw, affected scholars lose support from their peers, which are central to sustaining productivity. In this sense, professional distancing not only diminishes recognition of past work but also constrains the ability to generate new research, thereby offering an explanation for the observed decline in publications following speech sanctioning events.

6 Robustness

6.1 Within vs Outside Affected Scholar's Domain of Expertise

As a robustness check, we further probe into the underlying mechanism by distinguishing between work-related and non-work-related speech sanctioning events, categorising cases based on whether the controversial statement falls within or outside the scholar's domain of expertise. This distinction allows us to test competing explanations for the observed citation patterns. If peers were rationally updating their assessments of scholarly quality through Bayesian discounting, we would expect citation effects to be concentrated primarily on work-related controversies, since statements tied directly to a scholar's research domain provide legitimate grounds for reassessing the credibility of their academic contributions. However, if citation reductions extend beyond work-related speech sanctioning events, i.e. events ignited by statements unrelated to their academic work, this would provide stronger evidence that reputational concerns and social distancing are driving peer responses. As shown in Table 6, penalties are present in both types of cases. Our findings support the latter interpretation: citation penalties are present both within and outside the scholar's domain of expertise, providing further evidence for professional distancing as a mechanism.

Table 6: Mechanism: Inside and Outside a Scholar's Field of Expertise

	Inside Field	**Outside Field**
Post	0.0432***	-0.0048
	(0.0098)	(0.0094)
Treat × Post	-0.0593**	-0.0421*
	(0.0188)	(0.0176)
Fixed Effects:		
Article	Yes	Yes
Year	Yes	Yes
Journal × Publication Year	Yes	Yes
Model	Poisson	
S.E.: Clustered	Article	
Observations	173,246	178,002
Squared Correlation	0.79219	0.77223
Pseudo R^2	0.58101	0.57356
BIC	799,931.9	850,207.3

Notes: This table reports difference-in-differences estimates of how work-related (Column 1) and non-work-related (Column 2) speech sanctioning events affect citations to the focal scholar's prior work. Both specifications include article, year, and journal-by-publication-year fixed effects, and the standard errors are clustered at the article level. Significance levels: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\dagger}p < 0.1$.

6.2 'Active' Researchers

As an additional robustness check, we restrict the sample to scholars who remain 'research-active', defined as those with more than five publications both before and after the event. This ensures that we focus on individuals who continue to participate meaningfully in academic research rather than those who may be intentionally winding down or exiting the profession. As shown in Table 7, the citation and publication penalties persist within this research-active sample.

Table 7: Effect of Speech Sanctioning Events on the Annual Citations and Publications of 'Research-active' Scholars

	Publications	**Citations**
Post	0.0333	0.0281***
	(0.0719)	(0.0071)
Treat × Post	-0.2560**	-0.0398**
	(0.1109)	(0.0137)
Fixed Effects:		
Article	No	Yes
Author	Yes	No
Year	Yes	Yes
Journal × Publication Year	No	Yes
Model	Poisson	
S.E.: Clustered	Author & Treated Scholar	Article
Observations	2,642	300,693
Squared Correlation	0.52627	0.78223
Pseudo R^2	0.32636	0.57619
BIC	16,128.6	1,431,736.2

Notes: This table reports Poisson regression estimates of how speech sanctioning events affect annual research output. Column (1) uses the number of publications as the dependent variable and includes author and year fixed effects, with standard errors clustered by author and treated scholar. Column (2) uses citations as the dependent variable and includes article, year, and journal-by-publication-year fixed effects, with standard errors clustered by article. Significance levels: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\dagger}p < 0.10$.

7 Further Results

Although not central to our primary conceptual framework, we extend the analysis to examine whether speech sanctioning events generate spillovers onto the focal scholar's close collaborators. These are the colleagues who are most likely to share reputation, resources, and ongoing projects with the focal individual. From a theoretical perspective, this allows us to investigate whether the negative effects on the focal scholar spill over to their teams.

Leveraging the same difference-in-differences framework, we estimate how a collaborator's exposure to a speech sanctioning event influences their subsequent publication output. To this end, we focus on the scholars who are close coauthors of the affected scholars as of the event year and gather a sample of 82 coauthors associated with 471 similar control scholars. We define 'close coauthors' as the collaborators who produced the greatest number of joint publications with the focal scholar, with a minimum threshold of three coauthored papers in the ten years prior to the event. Table 8 reports the results of this regression, with 'close' co-authors publishing approximately 12.5% fewer publications than control group scholars.

Table 8: Impact of Treatment on the Yearly Publications of the Close Co-Authors

	Publications
Post	-0.0108
	(0.0378)
Treat $\times$ **Post**	-0.1361*
	(0.0641)
Fixed Effects:	
Author	Yes
Year	Yes
Model	Poisson
S.E.: Clustered	Author & Treated Scholar
Observations	6,103
Squared Correlation	0.66405
Pseudo R^2	0.46328
BIC	39,798.0

Notes: Estimates come from a Poisson regression of yearly publication counts among close co-authors. The coefficient of interest, *Treat* $\times$ *Post*, measures how co-authors' publication output changes after the focal scholar's speech sanctioning event relative to matched control co-authors. Standard errors are clustered at the author and treated-scholar level. Significance levels: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\dagger}p < 0.10$.

8 Limitations, Conclusion, and Discussion

8.1 Limitations

Our study has several limitations:

First, our study focuses on the United States, which limits the generalisability of our findings to other national contexts. This constraint primarily reflects data availability, as our dataset captures only events involving U.S.-based scholars. While we cannot empirically assess external validity, debates surrounding speech sanctioning events have also emerged in other countries (Norris, 2023; Scholars at Risk, 2023), suggesting that similar dynamics may operate elsewhere. We hope that future research can assemble comparable datasets across different geographic and cultural settings. Moreover, although our empirical setting is academia, the underlying mechanisms we document - peer responses to the speech sanctioning events, and subsequent productivity losses and changes in peer dynamics - are plausibly relevant in organisational contexts more broadly. Similar dynamics may unfold within firms, where professional networks, team interdependence, and reputational concerns likewise shape individual performance.

Second, our sample is limited to events that received sufficient media coverage to be included in the FIRE dataset. This necessarily introduces a bias toward more prominent cases. However, this focus also offers several empirical and theoretical advantages. First, empirically, speech sanctioning events are often internal and difficult to observe systematically; our dataset provides rare visibility into an otherwise opaque phenomenon. Comparable sampling strategies are common in prior studies of innovation (Azoulay et al., 2010; Widmann et al., 2025). Second, methodologically, visibility is essential for our identification strategy: peer responses can only be detected when information about the incident is sufficiently diffused for colleagues to be aware of it. Third, while our dataset captures relatively high-profile cases, these episodes provide insight into the mechanisms of sanctioning and professional distancing that are likely to arise even in less visible or lower-salience settings.

Third, our dataset likely contains incomplete coverage of speech sanctioning events, as no data collection effort can exhaustively capture all instances of speech sanctioning events in academia. However, FIRE's systematic approach to data collection - involving comprehensive searches of campus, local, and national media outlets, supplemented by cross-referencing with existing databases - represents the most rigorous attempt to date to document these phenomena. Crucially, we have no reason to believe that the data collection process introduces systematic biases that would affect our core findings.

Fourth, given the scale of this dataset, errors and misclassifications are inevitable. Wilson (2021) and Grady (2025) document several criticisms of the dataset's accuracy. For example, Wilson (2021) found several missing cases of scholars' sanctions. Nevertheless, Wilson (2021) explains, "To its credit,

FIRE updated its database to include the examples I found," and concludes, "Despite its flaws, the Scholars Under Fire database is chilling evidence of a massive number of attacks on professors in recent years." While FIRE's transparent methodology and open database reveal the dataset's limitations and incompleteness, we believe it remains reliable for the purpose of this study.

Fifth, our dataset encompasses a heterogeneous array of events spanning diverse topics, contexts, and severity levels. While we attempt to exploit this variation by, for example, distinguishing between statements made within versus outside scholars' domains of expertise, we lack sufficient statistical power to conduct subgroup analyses of specific speech categories. This heterogeneity reflects the inherent breadth of free speech as a concept, which by definition encompasses a wide spectrum of expression. As larger datasets become available, future research could provide more nuanced analyses of how different categories of speech generate varying professional impacts.

Finally, a potential critique of our study is that we do not evaluate the scientific validity or factual accuracy of the controversial statements in our dataset. While some events clearly involve false claims, we deliberately avoid adjudicating whether the statements are true or not. First, determining scientific or factual validity across diverse disciplines and contexts would introduce substantial researcher bias and methodological complexity. Second, our research question focuses on the professional consequences of being sanctioned for views, regardless of their accuracy - a distinction that is central to understanding academic freedom and its limitations. By maintaining this analytical neutrality, we can examine the phenomenon of speech sanctioning events without prejudging the merit of the underlying statements, thereby providing insights that are relevant regardless of one's position on the substantive issues involved.

8.2 Conclusion and Discussion

To conclude, this study examines the professional consequences of speech sanctioning events by analysing scholars who faced public backlash for their speech. Using a novel dataset from the Foundation for Individual Rights and Expression (FIRE) and employing a series of difference-in-differences designs, we find several results:

First, scholars involved in speech sanctioning events receive 4.8% fewer citations to their prior body of work relative to comparable papers published in the same journal, volume, and issue. Second, these scholars subsequently publish 20% fewer papers than their matched counterparts. Third, these effects extend beyond the focal scholar: close collaborators, those previously co-authoring with the focal individual, publish roughly 12.5% fewer papers after the event relative to comparable peers. Fourth, we observe evidence of professional distancing, as scholars with closer professional ties to the focal scholar reduce their citations disproportionately more than distant colleagues. Finally, we find that the citation penalty is present for cases ignited by events both inside and outside the scholar's area of

expertise, supporting a professional distancing mechanism rather than purely Bayesian updating.

Our findings have several implications for research on organisations and innovation.

First, we contribute to the literature on how knowledge is recognised and evaluated within organisations and professional communities. A substantial body of work shows that the credit accorded to ideas depends not only on their intrinsic quality but also on the social status, identity, and institutional affiliation of their creators. We extend this literature by demonstrating that speech sanctioning events can systematically reshape evaluative behaviour, altering how peers recognise existing work even when its underlying quality remains unchanged. The observed citation responses are consistent with a mechanism of professional distancing, whereby peers adjust their behaviour to manage reputational exposure. More broadly, similar peer-to-peer dynamics are likely to operate within firms and other organisational settings, where speech sanctioning events may influence how contributions are credited and how professional relationships evolve.

While our study focuses on the direct consequences of speech sanctioning events, the findings carry broader implications for professional discourse. The professional penalties we document create rational incentives for self-censorship among those who witness these outcomes. Although we cannot directly observe such anticipatory behaviour, the magnitude of the identified costs suggests that individuals may strategically moderate their public expression to avoid similar risks. In this way, speech sanctioning events may generate a broader chilling effect, whereby the fear of professional sanction discourages open discourse and constrains the exchange of ideas beyond the cases we directly observe. Future research could investigate the scope and persistence of these deterrent effects across different professional and organisational settings.

Second, we contribute to the literature by identifying a new form of disruption to the production of knowledge: speech sanctioning events. Prior research has shown that innovation can be impeded by shocks arising from scientific, institutional, or organisational disruptions. We extend this literature by showing that social and reputational conflict can have similar effects. Individuals involved in speech sanctioning events experience sustained declines in research output, highlighting how innovation depends not only on resources and incentives but also on the social environment in which work is embedded. This insight plausibly generalises to firms and other organisations in which productivity relies on collaboration, peer support, and reputational standing.

Third, we contribute to the literature on scandals by examining a phenomenon that shares similar structural features yet differs fundamentally in its origins. Scandals are typically defined as episodic, highly visible public events that trigger moral outrage and reputational consequences. Prior research shows that such events can disrupt careers, reshape collaboration networks, and alter organisational dynamics. Speech sanctioning events exhibit comparable visibility and reputational fallout but differ in a crucial respect: they do not stem from ethical or legal transgressions, but from the expression of

minority, contested, or unpopular views. By documenting comparable professional consequences in the absence of formal rule-breaking, our findings extend theories of scandals in organisational contexts.

Taken together, although our empirical setting is academia, the organisational dynamics we document plausibly extend to firms and other professional environments. Speech sanctioning events are not unique to universities: employees in corporate settings increasingly face public and internal scrutiny for expressing views perceived as controversial or misaligned with organisational norms. Related work on CEO activism shows that when corporate leaders take public stances on socio-political issues, their actions can meaningfully influence consumer behaviour, employee motivation, and firm value (Chatterji and Toffel, 2019; Mkrtchyan et al., 2024; Hou and Poliquin, 2024; Burbano, 2021; Mkrtchyan et al., 2024; Bondi et al., 2025). Our findings complement this literature by highlighting how speech sanctioning events can impose meaningful organisational costs through peer-to-peer behavioural responses and subsequent productivity declines. For managers, these results emphasise the importance of developing institutional processes that can anticipate and manage such events.

References

Adut, A. (2018). *Reign of appearances: The misery and splendor of the public sphere*. Cambridge University Press.

Ananthakrishnan, U. M., N. Basavaraj, S. R. Karmegam, A. Sen, and M. D. Smith (2025). Book bans in american libraries: Impact of politics on inclusive content consumption. *Marketing Science*.

Ayoubi, C., M. Pezzoni, and F. Visentin (2019). The important thing is not to win, it is to take part: What if scientists benefit from participating in research grant competitions? *Research Policy 48*(1), 84–97.

Azoulay, P., J. L. Furman, Krieger, and F. Murray (2015). Retractions. *Review of Economics and Statistics 97*(5), 1118–1136.

Azoulay, P., J. S. Graff Zivin, and J. Wang (2010). Superstar extinction. *The Quarterly Journal of Economics 125*(2), 549–589.

Azoulay, P., T. Stuart, and Y. Wang (2014). Matthew: Effect or fable? *Management Science 60*(1), 92–109.

Azoulay, P., J. M. Wahlen, and E. W. Zuckerman Sivan (2019). Death of the salesman but not the sales force: How interested promotion skews scientific valuation. *American Journal of Sociology 125*(3), 786–845.

Berger, J., A. T. Sorensen, and S. J. Rasmussen (2010). Positive effects of negative publicity: When negative reviews increase sales. *Marketing science 29*(5), 815–827.

Bikard, M., I. Fernandez-Mateo, and R. Mogra (2025). Standing on the shoulders of (male) giants: Gender inequality and the technological impact of scientific ideas. *Administrative Science Quarterly*.

Bikard, M., K. Vakili, and F. Teodoridis (2019). When collaboration bridges institutions: The impact of university–industry collaboration on academic productivity. *Organization Science 30*(2), 426–445.

Bol, T., M. De Vaan, and A. Van De Rijt (2018). The matthew effect in science funding. *Proceedings of the National Academy of Sciences 115*(19), 4887–4890.

Bondi, T., V. C. Burbano, and F. Dell'Acqua (2025). When (not) to talk politics in business: Experimental evidence. *Strategic Management Journal 46*(5), 1105–1119.

Burbano, V. C. (2021). The demotivating effects of communicating a social-political stance: Field experimental evidence from an online labor market platform. *Management Science 67*(2), 1004–1025.

Callaway, B. and P. H. Sant'Anna (2021). Difference-in-differences with multiple time periods. *Journal of econometrics 225*(2), 200–230.

Cambridge Dictionary (2025). Hate speech definition. `https://dictionary.cambridge.org/us/dictionary/english/hate-speech`.

Cavalli, G. (2022). How scientific organizations react to novel methodological advances: The impact of alphafold v1. *Working Paper*.

Chatterji, A. K. and M. W. Toffel (2019). Assessing the impact of ceo activism. *Organization & Environment 32*(2), 159–185.

Clark, C. J., L. Jussim, K. Frey, S. T. Stevens, M. Al-Gharbi, K. Aquino, J. M. Bailey, N. Barbaro, R. F. Baumeister, A. Bleske-Rechek, et al. (2023). Prosocial motives underlie scientific censorship by scientists: A perspective and research agenda. *Proceedings of the National Academy of Sciences 120*(48), e2301642120.

Cohen-Zada, D., I. Dayag, and N. Gershoni (2024). Effort peer effects in team production: Evidence from professional football. *Management Science 70*(4), 2355–2381.

Collis, M. R. (2025). Sexual misconduct and scientific production. *Working Paper*.

Deephouse, D. L. (2000). Media reputation as a strategic resource: An integration of mass communication and resource-based theories. *Journal of management 26*(6), 1091–1112.

Dumlao, J. M. and M. Teplitskiy (2025). Geographical diversity of peer reviewers shapes author success. *Proceedings of the National Academy of Sciences 122*(33), e2507394122.

Esposito, C. (2024). Knowledge spillovers within and across age groups of inventors. *Working Paper*.

Fombrun, C. and M. Shanley (1990). What's in a name? reputation building and corporate strategy. *Academy of management Journal 33*(2), 233–258.

Frey, K. and S. T. Stevens (2023). Scholars under fire: Attempts to sanction scholars from 2000 to 2022. Report, Foundation for Individual Rights and Expression.

Fry, C. and M. MacGarvie (2024). Author country of origin and attention on open science platforms: Evidence from covid-19 preprints. *Management Science 70*(8), 5426–5444.

Grady, M. (2025). The problem with fire's data. `https://www.chronicle.com/article/the-problem-with-fires-data`. The Chronicle of Higher Education.

Gulati, K. and C. Helmers (2025). Can political instability freeze science? evidence from the 2013 us government shutdown and scientific research in antarctica. *Working Paper*.

Harvard CAPS & Harris Poll (2021). Harvard caps/harris poll february 2021 crosstabs. Technical report, Harvard University Center for American Political Studies & Harris Poll.

Hegde, D. and B. Sampat (2015). Can private money buy public science? disease group lobbying and federal funding for biomedical research. *Management Science 61*(10), 2281–2298.

Honeycutt, N. and S. T. Stevens (2025). Sanctioned scholars: The price of speaking freely in today's academy. `https://www.thefire.org/research-learn/sanctioned-scholars-price-speaking-freely-todays-academy`.

Honeycutt, N., S. T. Stevens, and E. Kaufmann (2023). The academic mind in 2022: What faculty think about free expression and academic freedom on campus.

Hou, Y. and C. Poliquin (2024). Ceo activism and political mobilization. *Journal of Business Ethics*, 1–17.

Jaravel, X., N. Petkova, and A. Bell (2018). Team-specific capital and innovation. *American Economic Review 108*(4-5), 1034–1073.

Kessinger, R. (2024). *Speaking Up, Speaking Out, and Making Movements: How Employee Activists Raise Social, Political, and Moral Concerns at Work*. Ph. D. thesis, Massachusetts Institute of Technology.

Koffi, M. (2025). Innovative ideas and gender (in) equality. *American Economic Review 115*(7), 2207–2236.

Koffi, M. and M. Marx (2025). Cassatts in the attic: Is there a gender gap in the commercialization of science? *American Economic Journal: Applied Economics*.

Leahey, E. (2004). The role of status in evaluating research: The case of data editing. *Social Science Research 33*(3), 521–537.

Lu, S. F., G. Z. Jin, B. Uzzi, and B. Jones (2013). The retraction penalty: Evidence from the web of science. *Scientific reports 3*(1), 3146.

Luo, H. and L. Zhang (2022). Scandal, social movement, and change: Evidence from# metoo in hollywood. *Management Science 68*(2), 1278–1296.

Merton, R. K. (1973). *The sociology of science: Theoretical and empirical investigations*. University of Chicago press.

Mkrtchyan, A., J. Sandvik, and D. Xu (2024). Employee responses to ceo activism. *Journal of Accounting and Economics 78*(1), 101701.

Mkrtchyan, A., J. Sandvik, and V. Z. Zhu (2024). Ceo activism and firm value. *Management Science 70*(10), 6519–6549.

Mohnen, M. (2022). Stars and brokers: Knowledge spillovers among medical scientists. *Management Science 68*(4), 2513–2532.

Mumtaz, S. (2025). Lost in translation? science communication and the commercial diffusion of ideas. Working paper.

Norris, P. (2023). Cancel culture: Myth or reality? *Political studies 71*(1), 145–174.

Oettl, A. (2012). Reconceptualizing stars: Scientist helpfulness and peer performance. *Management science 58*(6), 1122–1140.

Petersen, A. M., E. M. Vincent, and A. L. Westerling (2019). Discrepancy in scientific authority and media visibility of climate change scientists and contrarians. *Nature communications 10*(1), 3502.

Phillips, D. P., E. J. Kanter, B. Bednarczyk, and P. L. Tastad (1991). Importance of the lay press in the transmission of medical knowledge to the scientific community. *New England journal of medicine 325*(16), 1180–1183.

Poege, F., F. Gaessler, K. Hoisl, D. Harhoff, and M. Dorner (2025). Filling the gap: The consequences of collaborator loss in corporate r&d. *Management Science*.

Pollock, T. G. and V. P. Rindova (2003). Media legitimation effects in the market for initial public offerings. *Academy of Management journal 46*(5), 631–642.

Priem, J., H. Piwowar, and R. Orr (2022). Openalex: A fully-open index of scholarly works, authors, venues, institutions, and concepts. *arXiv preprint arXiv:2205.01833*.

Qiu, J., Y. Chen, A. Cohn, and A. E. Roth (2024). Social media and job market success: A field experiment on twitter. *Working Paper*.

Rahmandad, H. and K. Vakili (2019). Explaining heterogeneity in the organization of scientific work. *Organization Science 30*(6), 1125–1145.

Reschke, B. P., P. Azoulay, and T. E. Stuart (2018). Status spillovers: The effect of status-conferring prizes on the allocation of attention. *Administrative Science Quarterly 63*(4), 819–847.

Resnik, D. B. (2008). Freedom of speech in government science. *Issues in science and technology 24*(2), 31.

Rindova, V. P., A. P. Petkova, and S. Kotha (2007). Standing out: How new firms in emerging markets build reputation. *Strategic organization 5*(1), 31–70.

Samila, S., A. Oettl, and S. Hasan (2022). Helpful behavior and the durability of collaborative ties. *Organization Science 33*(5), 1816–1836.

Scholars at Risk (2023). Free to think 2023. `https://www.scholarsatrisk.org/resources/free-to-think-2023/`.

Sofer, Y. (2024). Exploring the role of social media in the diffusion of economic research. *Working Paper*.

Stern, S. (2004). Do scientists pay to be scientists? *Management science 50*(6), 835–853.

Stuart, T. E. and W. W. Ding (2006). When do scientists become entrepreneurs? the social structural antecedents of commercial activity in the academic life sciences. *American journal of sociology 112*(1), 97–144.

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

Teodoridis, F., M. Bikard, and K. Vakili (2019). Creativity at the knowledge frontier: The impact of specialization in fast-and slow-paced domains. *Administrative Science Quarterly 64*(4), 894–927.

Teplitskiy, M., E. Duede, M. Menietti, and K. R. Lakhani (2022). How status of research papers affects the way they are read and cited. *Research policy 51*(4), 104484.

Tripodi, G., X. Zheng, Y. Qian, D. Murray, B. F. Jones, C. Ni, and D. Wang (2025). Tenure and research trajectories. *Proceedings of the National Academy of Sciences 122*(30), e2500322122.

Truffa, F. and A. Wong (2025). Undergraduate gender diversity and the direction of scientific research. *American Economic Review 115*(7), 2414–2448.

Vakili, K., F. Teodoridis, and M. Bikard (2022). Detrimental collaborations in creative work: Evidence from economics. *Organization Science 33*(5), 1741–1755.

Waldinger, F. (2012). Peer effects in science: Evidence from the dismissal of scientists in nazi germany. *The review of economic studies 79*(2), 838–861.

Widmann, R., M. E. Rose, and M. Chugunova (2025). Sexual misconduct, accused scientists, and their research. *Review of Economics and Statistics*, 1–29.

Wilson, J. (2021). Scholars under fire: Lessons from a database. `https://academeblog.org/2021/09/09/scholars-under-fire-lessons-from-a-database/`. Academe Blog.

Winkler, D., N. Wlömert, and J. Liaukonyte (2024). Separating the artist from the art: Social media boycotts, platform sanctions, and music consumption. *Cornell SC Johnson College of Business Research Paper*.

Wu, L., D. Wang, and J. A. Evans (2019). Large teams develop and small teams disrupt science and technology. *Nature 566*(7744), 378–382.

Yglesias, M. (2020). The real stakes in the david shor saga. `https://www.vox.com/2020/7/29/21340308/david-shor-omar-wasow-speech`. Vox.

Zheng, Y. and Q. Wang (2025). Blood diamonds? responses of open-source software developers to the facebook–cambridge analytica scandal. *Strategic Management Journal*.

Appendix

A Appendix: Descriptive Statistics

A.1 Plots of the Treated Scholars Over Time

The figures below plot the yearly count of treated scholars in each dataset, illustrating the timing of speech sanctioning events in our sample.

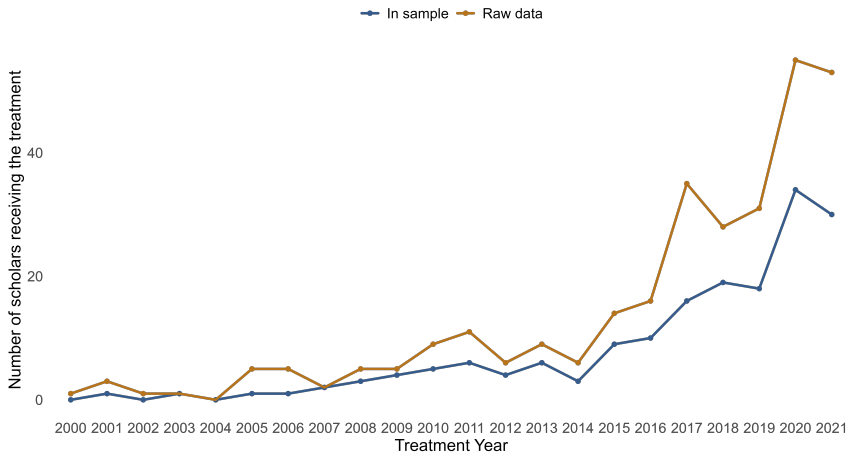


Figure A1: Number of treated scholars by year in the citation analysis.

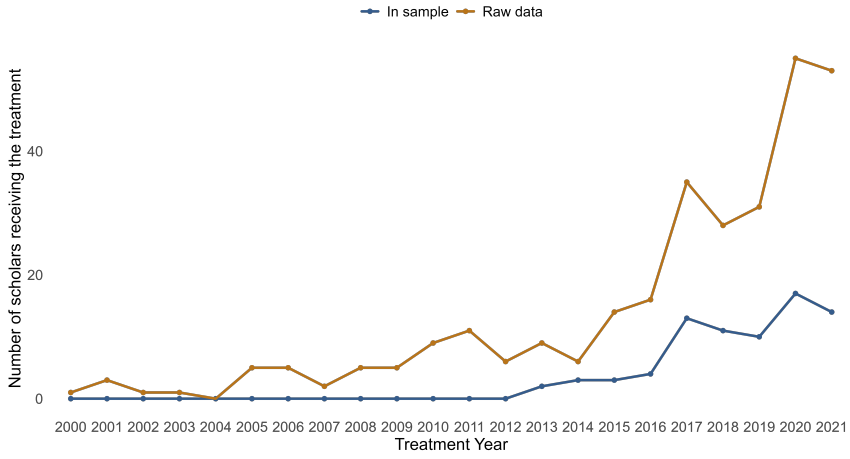


Figure A2: Number of treated scholars by year in the publication analysis.

A.2 Types of Works in the Citation Analysis

The table below shows the breakdown of works included in the citation analysis.

Type of work	Frequency
Article	334,141
Review	17,210

Table A1: Count of works by type

A.3 Descriptive Statistics on the Dependent Variables

The figures below plot the distributions of our primary dependent variables. To reduce the influence of extreme values, in our analysis, we exclude yearly citation counts above the 99th percentile (39 citations). The maximum value for citations before excluding extreme values is 2,577 citations per year. In the publication analysis, we also exclude yearly publications above the 99th percentile (33 publications). The maximum value for publications before excluding extreme values is 1,046 publications per year.

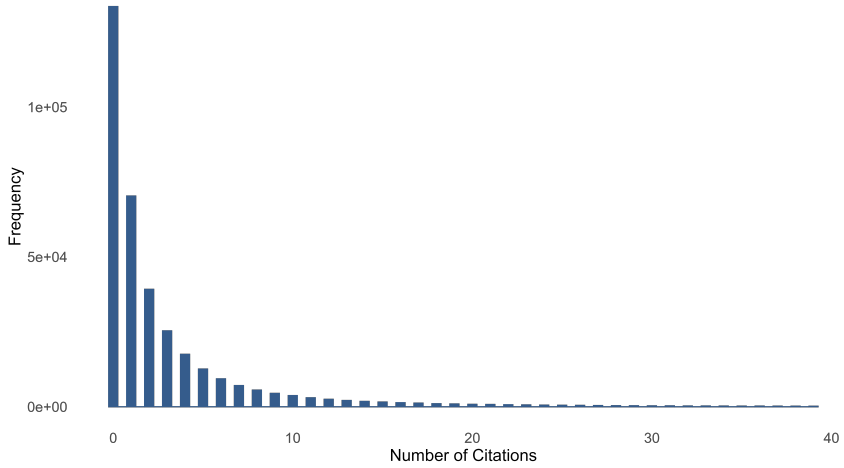


Figure A3: Distribution of the yearly number of citations.

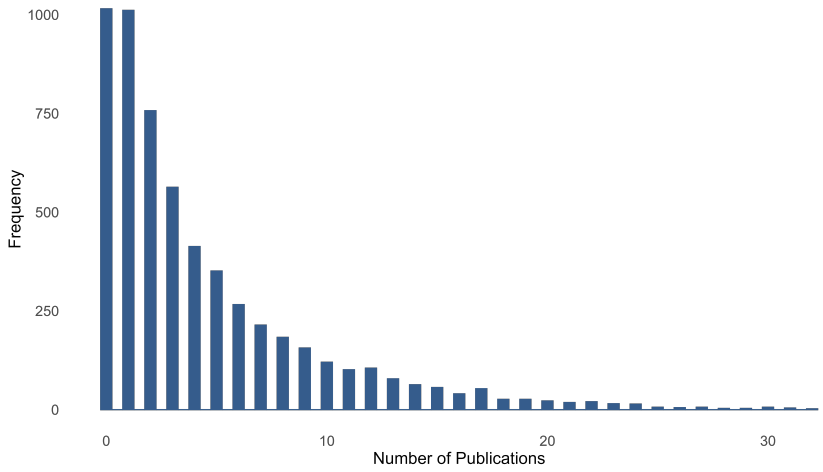


Figure A4: Distribution of the yearly number of publications.

A.4 Descriptive Statistics on the Close Co-authors of Treated Scholars

Table A2: Distribution of Coauthorship Intensity Between Treated Scholars and Their Closest Collaborators

Group	N	Mean	Median	Min	70th	80th	90th	95th	Max
Top coauthors	82	16.3	8	3	14	21	30	50	146

Notes: This table reports the distribution of collaboration intensity between treated scholars and their closest coauthors. "Top coauthor" refers to the top 2 most frequent coauthors of each treated scholar defined as having coauthored at least three papers in the 10 years before the event. *N* is the number of coauthors included in the analysis. Summary statistics include the mean, median, minimum, selected percentiles (70th, 80th, 90th, and 95th), and the maximum number of joint publications. All figures are aggregated at the scholar-coauthor dyad level.

B Appendix: Examples of Speech Sanctioning Events

Here we provide a handful of examples of cases captured in our dataset:

- **2013 — Milligan University (Theology).** A professor wrote an opinion article for *The Huffington Post*'s religion section about the marginal status of women in the Bible. He resigned.
- **2017 — Portland State University (Political Science).** A professor was the subject of a petition demanding the retraction of his peer-reviewed paper on the "supposed benefits of Western colonialism." The petition gathered 10,956 signatures.
- **2021 — University of California, Los Angeles (Medicine).** A professor was forced to resign as deputy editor of the *Journal of the American Medical Association (JAMA)* following outcry over a podcast in which he questioned structural racism.

- **2005 — University of Nevada, Las Vegas (Economics).** A professor attempted to illustrate time preferences by citing homosexuals as a group who tend to spend more readily because they are not thinking about starting a family. The initial backlash resolved with no sanction.
- **2012 — University of California, Los Angeles (Anthropology).** A professor came under scrutiny for posting a link to a site advocating for the cultural and academic boycott of Israel.

C Appendix: Inside and Outside Domain of Expertise Classification

We asked three research assistants to classify the events as either inside or outside the scholar's domain of expertise. To do so, the research assistants searched for information on each scholar to find their academic profile and compare this with the event.

All three research assistants independently adjudicated each event and made their classifications. Afterwards, any discrepancies where all three research assistants did not agree were resolved through discussion until consensus was reached.

D Appendix: Extreme Cases Classification

This section explains how we identified "extreme" cases from our analysis. We take two approaches, using research assistants and a large language model (GPT-4/4o), both asked to classify events using a clear definition of hate speech. The results are robust to both methods.

D.1 Extreme Cases: Labelled by Research Assistants

We asked three research assistants to classify statements as 'extreme'. To do so we provided them with the following definition of hate speech, "public speech that expresses hate or encourages violence towards a person or group based on something such as race, religion, sex, or sexual orientation." (Cambridge Dictionary). They also excluded cases involving statements classified as slurs or other forms of overtly offensive expression.

All three research assistants independently adjudicated each event and classified it as 'extreme' or not. Afterwards, any discrepancies where all three research assistants did not agree were resolved through discussion until consensus was reached.

D.2 Extreme Cases: Labelled by GPT4/4O

As a robustness check, we also leverage a large language model (GPT4/4O) to classify each statement as 'extreme' or not, providing it the same definition, "public speech that expresses hate or encourages violence towards a person or group based on something such as race, religion, sex, or sexual orientation." (Cambridge Dictionary).

Our results are robust and quantitatively similar using this alternative classification process, as shown in the two tables below:

Table A3: Robustness Check: Citation Analysis with GPT-Classification

Dependent Var.:	**Citations**
Post	0.0122†
	(0.0074)
Treat $\times$ **Post**	-0.0355*
	(0.0141)
Fixed Effects:	
Article	Yes
Year	Yes
Journal $\times$ Publication Year	Yes
Model	Poisson
S.E. Clustered	Article
Observations	265,581
Squared Cor.	0.78304
Pseudo R^2	0.58052
BIC	1,297,891.7

Notes: This table reports Poisson difference-in-differences estimates of how a speech sanctioning event affects annual citation counts. The dependent variable is the number of citations received by a paper in a given year. The key coefficient of interest, *Treat* $\times$ *Post*, captures the change in citations to papers authored by affected scholars after the event. All models include article, year, and journal-publication year fixed effects. Standard errors are clustered at the article level. The sample is restricted to papers published before the event to ensure that research quality is fixed prior to treatment. 11 scholars in our sample are excluded from this analysis. Significance levels: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\dagger}p < 0.1$.

Table A4: Robustness Check: Publication Analysis with GPT-Classification

Dependent Var.:	**Publication**
Post	0.0940
	(0.0580)
Treat × Post	-0.2586**
	(0.0901)
Fixed Effects:	
Author	Yes
Year	Yes
Model	Poisson
S.E. Clustered	Author & Treated Scholar
Observations	5,245
Squared Cor.	0.54263
Pseudo R^2	0.36093
BIC	29,934.3

Notes: This table reports Poisson difference-in-differences estimates of how speech sanctioning events affect a scholar's annual publication count. The dependent variable is the number of publications produced by each author in a given year. The coefficient of interest, *Treat* × *Post*, measures the change in publication output for affected scholars relative to matched controls after the event. All models include author and year fixed effects. Standard errors are clustered at the author and treated-scholar level. The sample is restricted to scholars who published both before and after the event, ensuring that the quality of research is fixed prior to treatment. This specification implements the GPT-classified approach as a robustness check. 3 scholars in our sample are excluded from this analysis. Significance levels: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, †$p < 0.1$.

E Appendix: Further Information on the Citation Analysis Data Construction

Starting from the 301 scholars identified in the Data section, we used the OpenAlex API to retrieve their complete publication records, yielding 36,569 works. Each scholar enters the sample with their first publication and exits with their last. To isolate the effect of speech sanctioning events, we restrict the sample to works published strictly before the treatment year. For example, if an event occurred in 2020, we include only publications from 2019 or earlier. This step leaves 28,025 works.

Following the approach in Azoulay et al. (2015), we compare pre-event works of affected scholars to a control group of papers published in the same journal, volume, and issue. This approach controls for the citation trajectories of the works and assumes that, absent a treatment, works published in the same journal, volume, and issue will follow on average similar citation trajectories. Due to missing values about the journal ID (4,506) and the issue and volume information (9,966), our preliminary dataset contains 13,553 works published by 294 treated scholars before the treatment.

We then extract control papers from OpenAlex, selecting up to ten per treated work from the same journal-volume-issue. Some treated works (261 of 13,553) lack eligible matches and are dropped, leaving 13,292 treated works linked to 127,367 control works. After obtaining yearly citation data, excluding works from scholars whose cases are followed by 'Turning point'[11], including in our sample

[11] Turning point often targets scholars in a certain year for events that have taken place in the past.

scholars involved in incidents until 2021 and works published up until 2020, excluding retracted works as well as works that are not in English, we gather a dataset of 259 treated scholars with 11,163 treated papers and 105,034 controls.

For the main analysis, we restrict to a balanced event window covering five years before and four years after treatment. Incorporating paper-level fixed effects automatically drops 45,869 works with zero citations in the observation window. We further require each treated paper to have at least three controls and trim the top 1% of the citation distribution to reduce outlier influence. At this point, we also exclude observations connected to scholars whose events have been tagged as extreme or scholars who have been eventually fired, terminated, or resigned following the events. The final dataset comprises 173 treated scholars with 5,182 treated papers and 35,076 control papers, totalling 351,248 observations.

F Appendix: Further Information on the Publication Analysis Data Construction

Starting from the 301 scholars identified in the Data section, we use the OpenAlex API to retrieve their complete publication histories and aggregate the information at the scholar-year level.

We then construct a pool of comparable scholars who have never been involved in any speech sanctioning events, matching them to treated scholars (as of the year before treatment) on: (1) primary research topics, (2) location, (3) number of publications, (4) number of citations, and (5) academic age. We limit the sample of treated scholars to the ones who received the treatment from 2012 onwards. This restriction has to be applied due to API restrictions.[12]

To calculate citations at the time of treatment, we: (i) download all publications for each treated scholar, (ii) subtract yearly citations after treatment from their total citation count, and (iii) sum citations received by each publication up to the treatment year. This allows us to calculate the total number of citations received by a scholar until the treatment. This reduces the treated sample from 301 to 253 scholars.

Next, we extract for each treated scholar (up to the treatment year) the total number of publications and citations, as well as their most frequent level-1 and level-2 research concepts. Because (1) all treated scholars were affiliated with a U.S. institution at the time of treatment, and (2) concept counts are calculated as the number of publications in which they appear, we can directly use these in our matching. We drop five treated scholars for whom concepts cannot be identified, leaving 248 in the dataset.

Once the matching criteria for the treated scholars are calculated, we can proceed to select the control group.

The control group is drawn from all U.S.-based researchers whose top level-1 and level-2 concepts include those of the treated scholar and whose total publications fall within $\pm 20\%$ of the treated

Considering these cases might cause issues for our analysis, 271,914 observations are dropped on 12,314 works divided into 1,204 and 11,144 treated and control works respectively with treatment years taking place mostly in 2016 and 2021 and affecting 23 treated scholars in total. The main results are still robust to their inclusion.

[12] OpenAlex provides yearly citation counts until 2012, so we can calculate the total number of citations received by each scholar's publication only until 2012. If a scholar was treated in 2005, for example, the calculations to identify potential control scholars would become too computationally intense.

scholar's count. From this pool, we select those whose total citations are within ±30% of the treated scholar's total. Using these thresholds, 182 treated scholars have at least one potential control.

However, for highly productive scholars, a ±20% publication gap or ±30% citation gap can still mean very large absolute differences (e.g., 328 publications and 14,338 citations). To improve comparability, we drop the top 1% of outliers in absolute differences, lowering the treated sample to 175.

We further match on academic age, imposing a maximum six-year difference in the year of first publication (the first quartile of the treated-control distribution). This leaves 138 treated scholars with at least one control. We also require that the difference in publications in the last pre-treatment year (T = -2) not exceed 10 to ensure comparable publication patterns. The final matched dataset contains 125 treated scholars and 1,470 control scholars.

Because some treated scholars have a very large number of matches, we randomly select up to 10 controls per treated scholar to avoid overweighting those with many potential matches. Once restricted to a time window that includes six periods before and four after the treatment year, excluding scholars targeted by Turning Point, and excluding observations connected to scholars whose events have been tagged as extreme or scholars who have been eventually fired, terminated, or resigned following the events, we gather a sample of 5,776 observations on 77 treated scholars and 466 control scholars.

G Appendix: Further Information on the Coauthor Analysis Data Construction

We construct a dataset of coauthors as follows. First, we restrict our sample to coauthors of treated scholars in the last ten years. We identify the top two coauthors for each treated scholar according to the number of papers coauthored in the last ten years.

Next, we identify comparable control scholars using the same procedure as in the publication-level analysis. Specifically, for each coauthor we gather, as of the treatment year, information on the number of publications, citations, country of main affiliation, year of first publication, and main research topics. We then select a pool of control scholars with similar characteristics: publications within ±20%, citations within ±30%, and a maximum difference of six years in the year of first publication.

We further refine the sample by (i) removing pairs with large absolute differences in publications or citations, (ii) retaining only scholars observed at least in the year before, the year of, and the year after the treatment, and (iii) keeping only coauthors with at least one matched control scholar. To conclude, we restrict the group of coauthors to those that have at least three publications with the focal scholars in the ten years before the event. The resulting dataset contains 82 coauthors matched to 471 control scholars.

H Appendix: Parallel Trends

This section presents the parallel-trends assessments underlying each of the main empirical specifications.

Table A5: Parallel Trends and Post-Period Effects: Citation Analysis

	Dynamic	**Post-Period Only**
Dependent Variable: Citations		
Treat × Period −5	0.0170	
	(0.0216)	
Treat × Period −4	0.0049	
	(0.0200)	
Treat × Period −3	0.0060	
	(0.0186)	
Treat × Period −2	0.0009	
	(0.0176)	
Treat × Period 0	-0.0257	
	(0.0168)	
Treat × Period 1	-0.0376*	
	(0.0179)	
Treat × Period 2	-0.0508*	
	(0.0198)	
Treat × Period 3	-0.0482*	
	(0.0231)	
Treat × Period 4	-0.0857**	
	(0.0286)	
Post		0.0235***
		(0.0067)
Treat × Post		-0.0491***
		(0.0128)
Fixed Effects:		
Article	Yes	Yes
Year	Yes	Yes
Journal × Publication Year	Yes	Yes
Model	Poisson	
S.E.: Clustered	Article	
Observations	351,248	351,248
Squared Cor.	0.78145	0.78145
Pseudo R^2	0.57704	0.57704
BIC	1,680,258.8	1,680,258.8

Notes: The dependent variable is the annual citation count to papers published before the speech sanctioning event. Column 1 reports the dynamic event-study coefficients, while Column 2 reports the baseline post-period difference-in-differences estimate. All specifications include article, year, and journal-by-publication-year fixed effects. Standard errors (in parentheses) are clustered at the article level. Significance levels: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\dagger}p < 0.10$.

Table A6: Parallel Trends and Post-Period Effects: Publication Analysis

	Dynamic	**Post-Period Only**
Dependent Var.: Publications		
Treat × Period -6	-0.1497	
	(0.1700)	
Treat × Period -5	-0.1110	
	(0.1618)	
Treat × Period -4	-0.1146	
	(0.1453)	
Treat × Period -3	-0.1010	
	(0.1595)	
Treat × Period -2	-0.1045	
	(0.1133)	
Treat × Period 0	-0.1557	
	(0.1371)	
Treat × Period 1	-0.3908**	
	(0.1387)	
Treat × Period 2	-0.4352**	
	(0.1386)	
Treat × Period 3	-0.2509*	
	(0.1324)	
Treat × Period 4	-0.1718	
	(0.1902)	
Post		0.0709
		(0.0584)
Treat × Post		-0.2201*
		(0.0967)
Fixed Effects:		
Author	Yes	Yes
Year	Yes	Yes
Model	Poisson	
S.E.: Clustered	Author & Treated Scholar	
Observations	5,776	5,776
Squared Cor.	0.55967	0.55413
Pseudo R^2	0.37383	0.37184
BIC	32,657.5	32,589.7

Notes: The dependent variable is the annual number of publications per scholar-year. The first column reports the dynamic event-study coefficients, while the second column reports the baseline post-period difference-in-differences estimate. All specifications include author and year fixed effects and are estimated with a Poisson model to account for the count nature of the dependent variable. Standard errors (in parentheses) are clustered at the author and treated scholar levels. Significance levels: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\dagger}p < 0.1$.

Table A7: Parallel Trends and Post-Period Effects: Coauthor Analysis

	Dynamic	**Post-Period Only**
Dependent Var.: Publications		
Treat × Period -6	0.1054	
	(0.1164)	
Treat × Period -5	-0.0316	
	(0.1330)	
Treat × Period -4	-0.0829	
	(0.1048)	
Treat × Period -3	-0.1159	
	(0.1034)	
Treat × Period -2	-0.0803	
	(0.1002)	
Treat × Period 0	-0.0128	
	(0.1145)	
Treat × Period 1	-0.2193†	
	(0.1171)	
Treat × Period 2	-0.2255*	
	(0.1079)	
Treat × Period 3	-0.2669**	
	(0.0964)	
Treat × Period 4	0.0975	
	(0.1244)	
Post		-0.0108
		(0.0378)
Treat × Post		-0.1361*
		(0.0641)
Fixed Effects:		
Author	Yes	Yes
Year	Yes	Yes
Model	Poisson	
S.E.: Clustered	Author & Coauthor Scholar	
Observations	6,103	6,103
Squared Cor.	0.66721	0.66405
Pseudo R^2	0.46500	0.46328
BIC	39,842.9	39,798.0

Notes: The dependent variable is the annual number of publications per coauthor-year for close collaborators (defined as the top 2 coauthors with at least three joint publications in the decade prior to the event). The first column reports the full dynamic event-study specification, while the second column reports the baseline post-period difference-in-differences estimate. All regressions include author and year fixed effects and are estimated using Poisson models. Standard errors (in parentheses) are clustered at the author and coauthor scholar levels. Significance levels: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\dagger}p < 0.1$.

I Appendix: Robustness to Alternative Estimators

I.1 Robustness to Alternative Estimators: Event Studies

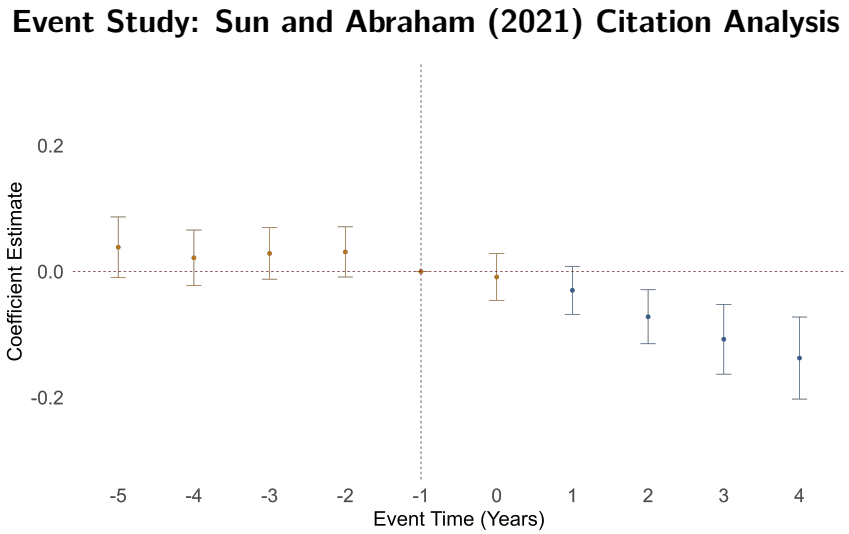


Figure A5: Estimates are obtained using the estimator proposed by Sun and Abraham (2021), which corrects for potential biases in staggered adoption designs. The dependent variable is the annual number of citations to papers published *prior* to the event, holding paper quality constant. Coefficients are relative to year -1, which is the omitted reference category. Vertical bars denote 95% confidence intervals; standard errors are clustered at the article level.

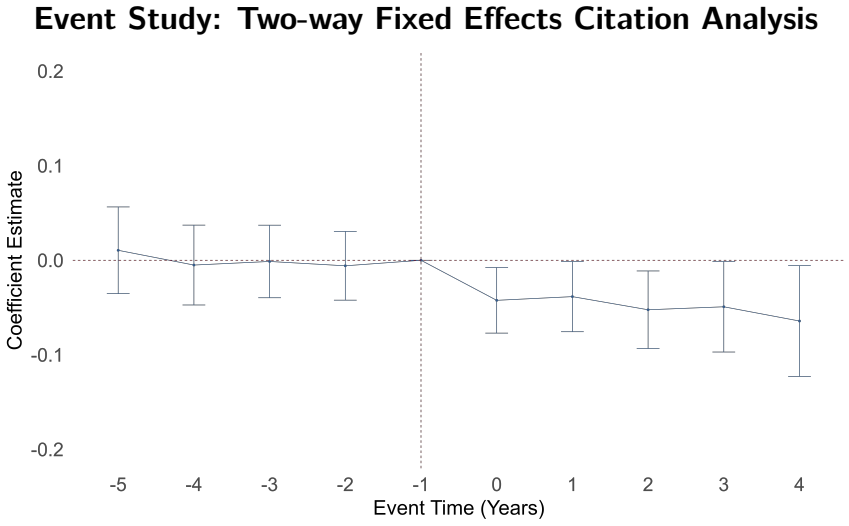


Figure A6: Estimates are obtained from a two-way fixed effects (TWFE) specification with article, year, and journal x publication year fixed effects. The dependent variable is the annual number of citations to papers published prior to the event, holding paper quality constant. Coefficients are relative to year -1, the omitted reference category. Vertical bars denote 95% confidence intervals, and standard errors are clustered at the article level.

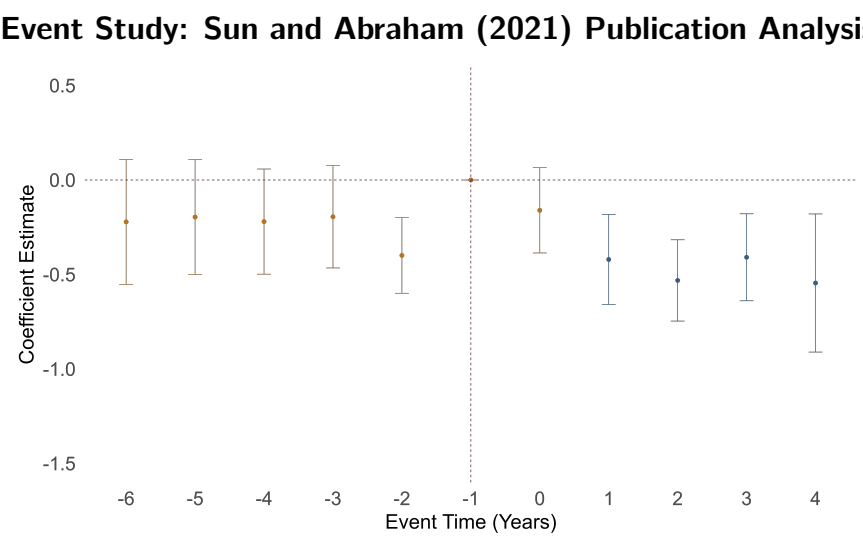


Figure A7: Estimates are obtained using the estimator proposed by Sun and Abraham (2021), which corrects for potential biases in staggered adoption designs. The dependent variable is the annual number of publications per scholar. Coefficients are relative to year -1, which is the omitted reference category. Vertical bars denote 95% confidence intervals; standard errors are clustered at the author and treated scholar level.

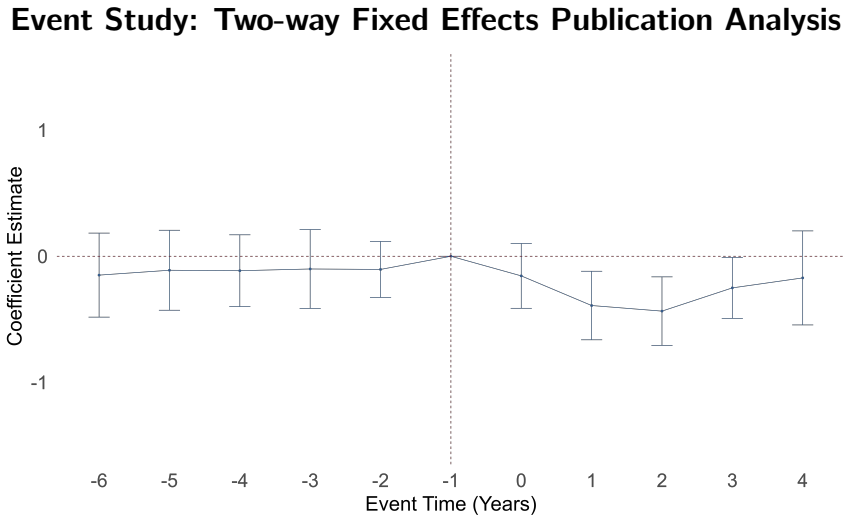


Figure A8: Estimates are obtained from a two-way fixed effects (TWFE) specification with author and year fixed effects. The dependent variable is the annual number of publications per scholar. Coefficients are relative to year -1, which is the omitted reference category. Vertical bars denote 95% confidence intervals; standard errors are clustered at the author and treated scholar level.

I.2 Robustness to Alternative Estimators: Difference-in-Differences

Table A8: Robustness Across Model Specifications: Citation Analysis

	Negative Binomial	**OLS**	**Sun and Abraham (ATT)**
Post	0.0309***	0.0484**	
	(0.0067)	(0.0186)	
Treat × Post	-0.0457***	-0.1455***	
	(0.0120)	(0.0373)	
ATT			-0.0551**
			(0.0175)
Fixed Effects:			
Article	Yes	Yes	Yes
Year	Yes	Yes	Yes
Journal × Publication Year	Yes	Yes	Yes
Model	Neg. Bin.	OLS	Poisson
S.E.: Clustered		Article	
Observations	351,248	351,351	351,248
Squared Cor.	0.77867	0.78049	0.78118
Pseudo R^2	0.27869	0.25087	7.0755
BIC	1,649,597.6	2,164,218.5	7.91e+10
Over-dispersion	7.6253	–	–

Notes: This table reports estimates from alternative model specifications evaluating the effect of speech sanctioning events on annual citation counts to pre-event publications. Columns 1 and 2 show results from Negative Binomial and OLS models, respectively. Column 3 reports the cohort-specific average treatment effect on the treated (ATT) using the Sun and Abraham (2021) estimator, implemented via a Poisson model. All specifications include article, year, and journal-by-publication-year fixed effects. Standard errors are clustered at the article level. Significance levels: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\dagger}p < 0.1$.

Table A9: Robustness Across Model Specifications: Publication Analysis

	Negative Binomial	**OLS**	**Sun and Abraham (ATT)**
Post	0.0659	0.1547	
	(0.0582)	(0.2457)	
Treat × Post	-0.1721*	-0.8239*	
	(0.0875)	(0.3908)	
ATT			-0.3869***
			(0.0995)
Fixed Effects:			
Author	Yes	Yes	Yes
Year	Yes	Yes	Yes
Model	Neg. Bin.	OLS	Poisson
S.E.: Clustered		Author & Treated Scholar	
Observations	5,776	5,776	5,776
Squared Cor.	0.54908	0.55171	0.56541
Pseudo R^2	0.15325	0.13186	1.8954
BIC	30,214.1	35,378.4	25,031,643.8
Over-dispersion	3.5505	–	–

Notes: This table reports estimates from alternative model specifications evaluating the effect of speech sanctioning events on annual publication counts. Columns 1 and 2 show results from Negative Binomial and OLS models, respectively. Column 3 reports the cohort-specific average treatment effect on the treated (ATT) using the Sun and Abraham (2021) estimator, implemented via a Poisson model. All specifications include author and year fixed effects. Standard errors are clustered at the author and treated scholar level. Significance levels: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\dagger}p < 0.1$.

J Appendix: Including Additional Cases

Table A10: Additional Samples (Dynamic): Citation Analysis

	Main	**Including 'Extreme'**	**Including FTR**	**Including 'Extreme' and FTR**
Dependent Var.: Citations				
Treat × Period −5	0.0170	0.0198	0.0171	0.0219
	(0.0216)	(0.0209)	(0.0200)	(0.0192)
Treat × Period −4	0.0049	0.0106	-0.0045	0.0025
	(0.0200)	(0.0193)	(0.0185)	(0.0178)
Treat × Period −3	0.0060	0.0112	0.0056	0.0122
	(0.0186)	(0.0180)	(0.0172)	(0.0166)
Treat × Period −2	0.0009	0.0072	-0.0078	-0.0036
	(0.0176)	(0.0173)	(0.0163)	(0.0159)
Treat × Period 0	-0.0257	-0.0209	-0.0253	-0.0235
	(0.0168)	(0.0164)	(0.0156)	(0.0152)
Treat × Period 1	-0.0376*	-0.0352*	-0.0295†	-0.0296†
	(0.0179)	(0.0173)	(0.0165)	(0.0158)
Treat × Period 2	-0.0508*	-0.0444*	-0.0443*	-0.0429*
	(0.0198)	(0.0192)	(0.0181)	(0.0175)
Treat × Period 3	-0.0482*	-0.0387†	-0.0336	-0.0315
	(0.0231)	(0.0224)	(0.0208)	(0.0200)
Treat × Period 4	-0.0857**	-0.0699*	-0.0731**	-0.0634*
	(0.0286)	(0.0272)	(0.0259)	(0.0247)
Fixed-Effects:				
Article	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes
Journal × Publication Year	Yes	Yes	Yes	Yes
Model			Poisson	
S.E.: Clustered			Article	
Observations	351,248	381,117	415,404	454,003
Squared Cor.	0.78145	0.77920	0.78172	0.77977
Pseudo R^2	0.57704	0.57749	0.57742	0.57807
BIC	1,680,258.8	1,817,616.5	1,979,216.1	2,160,926.0

Notes: This table reports period-specific treatment effects from Poisson regressions estimating the impact of speech sanctioning events on annual citations to pre-event publications across additional samples. All specifications include article, year, and journal × publication year fixed effects. Standard errors are clustered at the article level. The 'Main' sample is our preferred specification used in the main text. 'Extreme' refers to cases classified as involving extreme statements. 'FTR' refers to cases in which the scholar was fired, terminated, or resigned. Including 'Extreme' and FTR includes both types of cases. Significance levels: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\dagger}p < 0.10$.

Table A11: Additional Samples (Average Treatment Effect): Citation Analysis

	Main	**Including 'Extreme'**	**Including FTR**	**Including 'Extreme' and FTR**
Dependent Var.: Citations				
Post	0.0235***	0.0284***	0.0183**	0.0249***
	(0.0067)	(0.0065)	(0.0061)	(0.0060)
Treat × Post	-0.0491***	-0.0467***	-0.0369**	-0.0391***
	(0.0128)	(0.0125)	(0.0117)	(0.0114)
Fixed-Effects:				
Article	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes
Journal × Publication Year	Yes	Yes	Yes	Yes
Model			Poisson	
S.E.: Clustered			Article	
Observations	351,248	381,117	415,404	454,003
Squared Cor.	0.78122	0.77897	0.78154	0.77959
Pseudo R^2	0.57695	0.57741	0.57734	0.57800
BIC	1,680,280.9	1,817,649.5	1,979,235.9	2,160,978.9

Notes: This table reports post-period difference-in-differences estimates from Poisson regressions evaluating how speech sanctioning events affect annual citations to pre-event publications across alternative samples. All specifications include article, year, and journal × publication year fixed effects. Standard errors are clustered at the article level. The 'Main' sample is the preferred specification used in the main text. 'Extreme' refers to cases classified as involving extreme statements. 'FTR' refers to cases where the scholar was fired, terminated, or resigned. Including 'Extreme' and FTR includes both categories of cases. Significance levels: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\dagger}p < 0.10$.

Table A12: Additional Samples (Dynamic): Publication Analysis

	Main	**Including 'Extreme'**	**Including FTR**	**Including 'Extreme' and FTR**
Dependent Var.: Publications				
Treat × Period −6	-0.1497	-0.1093	-0.1966	-0.1745
	(0.1700)	(0.1598)	(0.1562)	(0.1416)
Treat × Period −5	-0.1110	-0.1076	-0.1368	-0.1432
	(0.1618)	(0.1528)	(0.1469)	(0.1342)
Treat × Period −4	-0.1146	-0.1582	-0.0442	-0.0617
	(0.1435)	(0.1447)	(0.1353)	(0.1382)
Treat × Period −3	-0.1010	-0.1220	-0.1496	-0.1614
	(0.1595)	(0.1539)	(0.1492)	(0.1393)
Treat × Period −2	-0.1045	-0.1348	-0.1484	-0.1800
	(0.1133)	(0.1092)	(0.1072)	(0.0975)
Treat × Period 0	-0.1557	-0.1356	-0.1403	-0.1281
	(0.1350)	(0.1284)	(0.1284)	(0.1172)
Treat × Period 1	-0.3908**	-0.3386**	-0.3326*	-0.2700*
	(0.1387)	(0.1314)	(0.1307)	(0.1144)
Treat × Period 2	-0.4352***	-0.4375***	-0.4851***	-0.4983***
	(0.1386)	(0.1323)	(0.1310)	(0.1195)
Treat × Period 3	-0.2509*	-0.1956	-0.3450**	-0.2768*
	(0.1234)	(0.1242)	(0.1204)	(0.1138)
Treat × Period 4	-0.1718	-0.1493	-0.3010	-0.2074
	(0.1902)	(0.1730)	(0.1840)	(0.1551)
Fixed-Effects:				
Author	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes
Model		Poisson		
S.E.: Clustered		Author & Treated Scholar		
Observations	5,776	6,554	6,783	7,900
Squared Cor.	0.55967	0.54139	0.54232	0.53229
Pseudo R^2	0.37383	0.36248	0.35894	0.35188
BIC	32,657.5	37,138.1	38,587.1	45,332.1

Notes: This table reports event-study coefficients for publication output across multiple sample definitions. All models are estimated using Poisson regressions with author and year fixed effects. Standard errors are clustered at the author and treated scholar level. The 'Main' sample is our preferred specification, which is used in the main text. 'Extreme' refers to cases classified as involving extreme statements. 'FTR' refers to cases where the scholar was fired, terminated, or resigned. Including 'Extreme' and FTR includes both types of cases. Significance levels: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\dagger}p < 0.10$.

Table A13: Additional Samples (Average Treatment Effects): Publication Analysis

	Main	**Including 'Extreme'**	**Including FTR**	**Including 'Extreme' and FTR**
Dependent Var.: Publications				
Post	0.0709	0.0793	0.0702	0.0735
	(0.0584)	(0.0523)	(0.0531)	(0.0449)
Treat × Post	-0.2201*	-0.1817*	-0.2582**	-0.2004*
	(0.0967)	(0.0923)	(0.0877)	(0.0814)
Fixed Effects				
Author	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes
Model		Poisson		
S.E. Clustered		Author & Treated Scholar		
Observations	5,776	6,554	6,783	7,900
Squared Cor.	0.55413	0.53717	0.53760	0.52844
Pseudo R^2	0.37184	0.36099	0.35690	0.35031
BIC	32,589.7	37,053.4	38,532.3	45,263.7

Notes: This table reports average treatment effects (ATEs) from Poisson regressions estimating post-event changes in annual publication output across additional samples. All specifications include author and year fixed effects. Standard errors are clustered by author and treated scholar. The 'Main' sample is our preferred specification. 'Extreme' refers to cases involving extreme statements. 'FTR' refers to cases where the scholar was fired, terminated, or resigned. Including 'Extreme' and FTR includes both types of cases. Significance levels: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\dagger}p < 0.10$.

K Appendix: Excluding Self Citations

Table A14: Event-Study and Post-Period Effects on Citations Excluding Self Citations

	Dynamic	**Post-Period Only**
Dependent Var.: Citations		
Treat × Period −5	0.0006	
	(0.0219)	
Treat × Period −4	-0.0069	
	(0.0202)	
Treat × Period −3	-0.0074	
	(0.0189)	
Treat × Period −2	-0.0024	
	(0.0179)	
Treat × Period 0	-0.0250	
	(0.0170)	
Treat × Period 1	-0.0303†	
	(0.0182)	
Treat × Period 2	-0.0417*	
	(0.0200)	
Treat × Period 3	-0.0347	
	(0.0234)	
Treat × Period 4	-0.0717*	
	(0.0290)	
Post		0.0207**
		(0.0067)
Treat × Post		-0.0324*
		(0.0130)
Fixed Effects:		
Article	Yes	Yes
Year	Yes	Yes
Journal × Publication Year	Yes	Yes
Model	Poisson	
S.E.: Clustered	Article	
Observations	350,676	350,676
Squared Cor.	0.78151	0.78128
Pseudo R^2	0.57709	0.57701
BIC	1,675,981.1	1,675,993.5

Notes: The dependent variable is yearly citations excluding self-citations for articles in the analysis sample. The first column reports coefficients from the dynamic event-study specification, while the second column reports the corresponding post-period difference-in-differences estimates. All regressions include article, year, and journal × publication year fixed effects. Standard errors (in parentheses) are clustered at the article level. Significance levels: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\dagger}p < 0.1$.

L Data Availability and Generative AI Disclosure

Data Availability Statement: The data that supports the findings of this study is fully open source and available at `https://www.thefire.org/research-learn/scholars-under-fire` and at `https://openalex.org/`. The exact data used in our analysis, as well as a full replication package, will be made available on a public repository by the authors upon acceptance for publication.

Generative AI Disclosure: During the preparation of this work, the authors used large language models to harmonise text and to categorise the speech sanctioning events in a robustness check. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.